

University of Pennsylvania Carey Law School

Penn Law: Legal Scholarship Repository

Faculty Scholarship at Penn Law

10-19-2019

The Effect of Police Oversight on Crime and Allegations of Misconduct: Evidence from Chicago

Bocar A. Ba

University of Pennsylvania Carey Law School

Roman G. Rivera

Columbia University

Follow this and additional works at: https://scholarship.law.upenn.edu/faculty_scholarship



Part of the [Behavioral Economics Commons](#), [Criminology and Criminal Justice Commons](#), [Law and Economics Commons](#), [Law Enforcement and Corrections Commons](#), [Policy Design, Analysis, and Evaluation Commons](#), [Public Affairs Commons](#), [Public Economics Commons](#), and the [Public Policy Commons](#)

Repository Citation

Ba, Bocar A. and Rivera, Roman G., "The Effect of Police Oversight on Crime and Allegations of Misconduct: Evidence from Chicago" (2019). *Faculty Scholarship at Penn Law*. 2109.
https://scholarship.law.upenn.edu/faculty_scholarship/2109

This Article is brought to you for free and open access by Penn Law: Legal Scholarship Repository. It has been accepted for inclusion in Faculty Scholarship at Penn Law by an authorized administrator of Penn Law: Legal Scholarship Repository. For more information, please contact PennlawIR@law.upenn.edu.

The Effect of Police Oversight on Crime and Allegations of Misconduct: Evidence from Chicago

Roman G. Rivera[‡]

Bocar A. Ba[†]

October 19, 2019

Abstract

Does policing the police increase crime? We avoid simultaneity effects of increased public oversight during a major scandal by identifying events in Chicago that only impacted officers' self-imposed monitoring. We estimate crime's response to self- and public-monitoring using regression discontinuity and generalized synthetic control methods. Self-monitoring, triggered by police union memos, significantly reduced serious complaints without impacting crime or effort. However, after a scandal, both civilian complaints and crime rates rise, suggesting that higher crime rates following heightened oversight results from de-policing and civilian behavior simultaneously changing. Our research suggests that proactive internal accountability improves police-community relations without increasing crime.

Many thanks to Robert LaLonde for his guidance and support. Special thanks go to Scott Ashworth, Pat Bayer, Christian Hansen, Steve Levitt, John MacDonald, and Arnaud Maurel for their help throughout this project. We are grateful to Amani Abou Harb, Joe Altonji, Arjada Bardhi, Amy Claessens, Rob Garlick, Felipe Goncalves, Justin Holz, Matthew Johnson, Dan Kowalski, Arvind Krishnamurthy, Matt Masten, Marjorie McElroy, Steve Mello, Magne Mogstad, Harry Pei, Nicola Persico, Guillaume Pouliot, Canice Prendergast, Nayoung Rim, Seth Sanders, Modibo Sidibe, Juan Carlos Suárez Serrato, Bruno Strulovici, Huseyin Yildirim, and seminar participants at Duke and Chicago. We thank Sam Stecklow, the Invisible Institute, Craig Futterman, Jamie Kalven, and Jacob Kaplan for help with the data. Emma Herman and Jenny Jiao provided outstanding research assistance. We thank the Duke Economics Department for generous financial support. Any errors are our own. Contact: [‡]rgr2122@columbia.edu, [†] bocarba@law.upenn.edu.

1 Introduction

Misconduct scandals have rocked multiple fields, including finance, entertainment, academia, medicine, and, most notably, policing.¹ Opponents of increased oversight have claimed that any form of increased monitoring may do more harm than good by decreasing productivity. This discussion has been particularly relevant in policing: theories of “de-policing” (MacDonald [2019], Owens [2019]) or the “Ferguson effect” (Mac Donald [2016]) propose that increased oversight will drive officers to reduce effort, resulting in higher crime rates. Yet, empirical evidence supporting such theories are often dependent on public scandals to induce shocks to oversight (Prendergast [2001], Shi [2009], Heaton [2010], Cassell [2018]). This paper provides new evidence on the effect of oversight on crime and police effort in a scandal-free environment.

To evaluate the effect of oversight on officer conduct, effort, and crime, we use detailed misconduct and crime data from the Chicago Police Department (CPD) and the Federal Bureau of Investigation’s Uniform Crime Reporting (UCR) from 2007 to 2016. We distinguish between two types of oversight: self-monitoring by officers results from an increase in their perceived cost of receiving misconduct accusations; public-monitoring by civilians results from an increase in their perceived benefit of alleging misconduct. By identifying events that solely increase self-monitoring, we are able to isolate the effect of oversight on our outcomes of interest and avoid the simultaneity bias that accompanies increased public monitoring during scandals. We find that an increase in self-monitoring significantly reduces allegations of constitutional violations, but it does not reduce officer effectiveness, as neither arrests nor crime are affected.

We contrast the self-monitoring results with estimates of the effect of a major policing scandal in Chicago on misconduct, crime, and effort. We find evidence confirming “de-policing” during scandal periods, as rates of murder and robbery increase without corresponding increases in arrest rates. Since major scandals are characteristically highly public and impact civilian opinion, values, and incentives regarding law enforcement, such scandals simultaneously affect both civilian and officer behavior. Incentives for civil-

¹For example, see finance (Keefe [2014], Stevenson and Goldstein [2016]), entertainment (Moniuszko [2017], Kantor and Twohey [2019]), medicine (Rabin [2015]), academia (Binder et al. [2016], Thompson and Clark [2019]).

ians to commit crime increase due to civil disorder, mistrust of officers, and decreased civilian cooperation, holding constant officer effort and enforcement. We term this “simultaneity bias”, and we provide evidence consistent with its existence: in the aftermath of a scandal, civilian complaints, from both victims (filing service focused complaints)² and suspects (filing complaints alleging constitutional violations) of crime, significantly increase despite officers simultaneously decreasing enforcement.

Using scandals³ as sources of exogenous variation is intuitive due to the noticeable and direct effects they have on their population of interest. However, by their nature, scandals temporarily and drastically change the incentives of both the principal (an officer suspected of misconduct) and agent (a civilian alleging misconduct), engendering a simultaneity bias in estimates of the effect of increased oversight, decreased trust, and heightened sensitivity and dissatisfaction with the police and institutions on outcomes of interest (Lee and Suen [Forthcoming]). To isolate the impact of oversight alone, it is necessary to identify events which solely impact officers’ incentives, as oversight agencies generally do not rely on heightened vigilance from all civilians.

To provide new evidence on the impact of oversight on misconduct, crime, and arrests, we rely on the assumption that our first two events only impacted officers and increased their self-monitoring. These events are articles in the monthly newsletter published by the city’s police union whose sole audience is Chicago police officers. The first article, the “FOP Note”, was published in 2009, and it notified officers of the risks of receiving complaints.⁴ The second article, the “Kalven Note”, was published in 2014, and it notified officers that complaints filed against them would be made public in the future. In contrast, the third event is the scandal that erupted in late 2015 following the public release of footage of the officer-involved-shooting (OIS) of Laquan McDonald. The articles have the advantage of being published in a periodical only read by officers, in a period free of major scandals involving the CPD, and, in the case of the first event, before the public

²Throughout this paper, we use the terms “failure to provide service,” or “service focused,” complaints involving individuals who sought help from the police. We also use the terms “constitutional violations,” or “serious,” complaints involving individuals engaged by the police on suspicion of criminal activity.

³For example, highly publicized cases of police brutality, corruption, discrimination, or sexual harassment.

⁴Unlike Mac Donald [2016], this event occurred in a period free of major scandal and pre-national debate on reforming the police.

debate about police accountability erupted across the United States (see Fig. 2). Our design isolates the effect of self-imposed monitoring by police officers from the impact of changes in oversight following a period of public scrutiny.

We take advantage of Chicago micro-level data to capture crime and officers' behavioral responses to oversight. Similar to [Anderson \[2014\]](#), we estimate the impact of oversight on complaints, crimes, and arrests using a regression discontinuity design (RD). We find that the number of complaints decreases after the release of the notes but increases after the OIS Scandal. The abrupt change in the number of constitutional violation complaints after each event suggests that officers altered their behavior with suspects. More precisely, we find that the increase in self-monitoring decreases those complaints by 22-35% ($p < 0.01$) relative to the baseline; in contrast, after the scandal, complaints increase by more than 61% ($p < 0.01$). We do not find evidence that the other types of complaints are affected by the change in self-monitoring, suggesting that officers did not adjust their behavior when interacting with non-suspects or colleagues (workplace complaints). However, both service-related and constitutional violation complaints significantly increase when there is an increase in public monitoring triggered by the scandal. The simultaneity bias associated with the scandal makes it difficult to identify how the different types of oversight impact officers' conduct.

The analysis indicates that there was no discontinuous change in crime rates around the increase in self-monitoring. In combination with the finding that the notes decreased constitutional violations, this suggests that officers improved their treatment of suspects without impacting public safety. On the other hand, there is a sharp increase in crime around the OIS Scandal. The total number of crimes in each beat increased by 0.64 per 10,000 residents (5.4% relative to baseline mean).

Following [Mas \[2006\]](#), we use arrests as a measure of officer effort. The first note sent by the union increased non-property crime-related arrests when using a local linear RD, but the results are neither statistically nor economically significant. We find no evidence that officers changed their arrest patterns when the Kalven Note was released. From these findings, we conclude that self-monitoring did not induce officers to decrease effort. In contrast, the OIS Scandal does alter arrest patterns, as crime increased with no corresponding change in arrests; however, as argued previously, it is difficult to interpret this

result due to the simultaneity bias.

Finally, using generalized synthetic control methods (GSC) from [Xu \[2017\]](#), we evaluate the effect of these events on crime rates in Chicago relative to other major U.S. cities. In the spirit of the synthetic control method from [Abadie and Gardeazabal \[2003\]](#), [Abadie et al. \[2010, 2015\]](#), this approach combines the interactive fixed effect models from [Bai \[2009\]](#) with a cross-validation method to build a counterfactual for Chicago. This counterfactual Chicago is constructed from the police departments of other major U.S. cities. In our preferred specification, the control group includes only cities with at least 500,000 residents that were never subject to a Department of Justice (DOJ) investigation⁵, i.e. subject to public monitoring.

Using GSC, we do not find evidence that crime increases when self-monitoring increases. We do find evidence that rates of property crime, murder, and robbery increase significantly after the scandal. Our preferred specification indicates that the OIS Scandal is associated with 18.4 additional murders in Chicago per month ($p < 0.01$) in the year following the scandal. For comparison, self-monitoring is associated with fewer than two extra murders per month, but this is not statistically significant from zero.

Our analysis that compares Chicago to other U.S. cities is robust to multiple specifications. Using the synthetic control method from [Abadie et al. \[2010, 2015\]](#), we find similar results to the GSC method. Our results are also robust to the inclusion of cities with smaller populations (population of at least 250,000) or subject to a DOJ investigation.

[Lee and Suen \[Forthcoming\]](#) provide theoretical explanations of the mechanisms behind our empirical findings during a scandal. We find evidence that alleged victims are strategic in reporting misconduct.⁶ Our findings suggests that civilian behavior toward the police is not exogenous during volatile periods: in periods of high scrutiny of the police, the number of complaints against the police increases significantly, despite evidence of less aggressive enforcement. These simultaneous responses to the scandal from both civilians and officers make it difficult to draw causal conclusions on oversight's impact

⁵Investigation of police departments from the DOJ are often triggered by scandals. Therefore we view the DOJ as part of public-monitoring.

⁶Post-scandal, victims recognize that the police department is more likely to respond to police complaints and that they have the support from various groups (ACLU, DOJ, etc.), so are more willing to come forward.

on crime using scandals.

Moreover, [Pei and Strulovici \[2019\]](#) provide theoretical explanations of the mechanisms behind civilian and officer behavior. Applying their findings to police officers, the authors show that when a principal (police) is accused of multiple instances (types) of misconduct, which would possibly result in a large punishment, her decisions to commit misconduct against different types of agents (civilians) are substitutes. Additionally, they show that agents that have been separately victimized report in complementary fashion. We show that an increase in police self-monitoring results in a decrease in complaints mostly from suspects of crimes while service related complaints increase marginally. This suggests that police officers may strategically abuse only one type of complainant at the margin and that different targets of bad behavior are substitutes. Additionally, after the scandal, both suspects and victims of crime increased their number of complaints; this suggests complementary responses of different agents to an increase in public scrutiny specifically regarding excessive use of force.

This paper contributes to several strands of the literature: (1) the deterrent effect of police on crime, (2) the impact of police oversight on crime, and (3) the objective function of police.

The deterrent effect of police on crime Since [Levitt \[1997\]](#), researchers have found evidence suggesting crime responds to police staffing levels ([Evans and Owens \[2007\]](#), [Fu and Wolpin \[2017\]](#), [Chalfin and McCrary \[2018\]](#), [Mello \[2019\]](#)), deployment and tactics ([Di Tella and Schargrodsky \[2004\]](#), [Draca et al. \[2011\]](#), [Klick and Tabarrok \[2005\]](#)⁷), and problem-oriented policing ([Miller and Segal \[2018\]](#), [Owens et al. \[2018\]](#)).

However, as argued in [Becker and Stigler \[1974\]](#), there is no reason to assume that officers never engage in malfeasance while enforcing the law. Hence, officer misconduct could impact multiple dimensions of policing: suspensions or re-assignments may reduce manpower; deployment and tactics may be misused due to overly aggressive policing or racial profiling; or the efficacy of problem-oriented policing may be diminished by damaged community relations. Our paper contributes to this literature by measuring the

⁷Also see [Blanes i Vidal and Kirchmaier \[2017\]](#), [Heaton, Hunt, MacDonald, and Saunders \[2016\]](#), [Chalfin and McCrary \[2018\]](#), [Weisburd \[2018\]](#), [i Vidal and Mastrobuoni \[2018\]](#)

impact of changes in different forms of oversight on relevant outcomes such as crime, arrests, and misconduct allegations.

The impact of police oversight on crime A series of articles estimating the effect of oversight on officer conduct have used scandals in quasi-experimental designs that generated increased media attention and judicial scrutiny; they find increasing oversight decreases police effort and increases crime rates ([Prendergast \[2001\]](#), [Shi \[2009\]](#), [Heaton \[2010\]](#)).

We argue that one cannot interpret such estimates as the effect of oversight on policing because scandals induce simultaneous and significant changes in civilian incentives which impact crime rates. This is informed by empirical and theoretical evidence. For example, [Desmond et al. \[2016\]](#) documents that high-profile cases of police violence result in citizens being less likely to report crimes, and this can increase crime opportunities independent of police effort. Furthermore, [Persico \[2002\]](#) suggests that unfair policing of a population will encourage crime and provides theoretical conditions under which more equitable policing would actually decrease crime.

We extend this literature in two ways: we find further evidence that scandals increase crime, and we provide evidence that civilian and officer incentives simultaneously change during a scandal. After the OIS Scandal, complaints about both officer conduct and constitutional violations as well as murders and robberies increase significantly, despite arrest rates remaining unchanged. In contrast, we find that events which only increase self-monitoring significantly decrease complaints of constitutional violations and have no impact on crime rates or arrests.

A related literature shows that officers are highly responsive to managerial directives ([Mas \[2006\]](#)). In particular, studies found that more internal monitoring leads to higher hit rates for stop-and-frisk in New York ([Mummolo \[2018\]](#)) and fewer violent crimes in New Orleans ([Cheng and Long \[2018\]](#)). Our results support these findings, as we find police union directives that increased self-monitoring among police officers strongly impacted officer conduct, resulting in fewer allegations of constitutional violations and no increase in crime or decrease in arrests.

The objective function of police Our paper contributes to the economic literature on the objective function of police officers—a subject that is highly relevant for policymakers. As argued by [Manski \[2006\]](#), social planners only have partial knowledge of how policies affect populations. Furthermore, there is ample theoretical literature on the cost of enforcement against innocent individuals (e.g. false arrests or failures to provide service) by [Donohue and Levitt \[2001\]](#), [Persico \[2002\]](#), [Manski and Nagin \[2017\]](#), which contributes to understanding how officers may engage in misconduct as they enforce the law ([Becker and Stigler \[1974\]](#), [Benoit and Dubra \[2004\]](#)). However, there is little empirical evidence on how officer misconduct impacts crime, and most of the literature on the costs and benefits of proactive policing has focused on racial differences ([Persico \[2002\]](#), [Durlauf \[2006\]](#), [Manski \[2005, 2006\]](#)).⁸

In this vein, we provide two main contributions. First, we find that an increase in self-monitoring reduces misconduct without impacting the level of crime, while scandals have a detrimental impact on the level of violence. Our results on public-monitoring are in line with the theoretical results of [Prat \[2005\]](#): an agent (officer) who knows that his action is observed by the principal (public) has an incentive to behave in a conformist manner thus permitting more crime, which can hurt the principal (public)—this is also in line with theories of “de-policing.”

Secondly, we show that an increase in self-monitoring results in a reduction of allegations of constitutional violations, which are primarily filed by criminal suspects rather than by victims of crime. This result complements the findings of [Ba \[2018\]](#) who shows that a policy that would reduce the cost of filing a complaint would mainly impact crime victims, rather than suspects.

The rest of the paper is organized as follows. Section 2 provides background information on police oversight in Chicago. Section 3 presents the effect of oversight on the number of complaints filed against officers, crime, and arrest using micro-level data in Chicago. Section 4 discusses the effect of oversight on crime comparing Chicago to other

⁸There is a large literature providing compelling empirical research designs to test for the presence of racial bias in motor vehicle stops and searches ([Knowles et al. \[2001\]](#), [Anwar and Fang \[2006\]](#), [Antonovics and Knight \[2009\]](#)). However, those designs do “not allow for the possibility of false accusation by police or planting of evidence” (See footnote 5 in [Knowles et al. \[2001\]](#)). [Dharmapala and Miceli \[2012\]](#) provide legal and theoretical explanations of officer incentives to make searches, seizures, and false arrests.

major U.S. cities. Section 5 explores the robustness of our results. Section 6 provides a discussion on the potential policy implications and channels that explain our results. Section 7 concludes.

2 Background

The past ten years have witnessed significant fluctuations in police oversight in Chicago. Figure 1 documents these fluctuations, where we take civilian complaints to be our main metric of oversight between January 2007 and May 2016. This figure presents the total number of monthly allegations of misconduct against Chicago police officers from January 2007 to May 2016.⁹

We divide Figure 1 into four periods, segmented by the three previously identified events. From January 2007 to October 2009, the average number of complaints is around 541 per month. The second period, November 2009 to July 2014, has an average of 361 complaints per month; this corresponds to decline of 34% relative to the first period. The third period, August 2014 to October 2015, has an average of 169 complaints per month, which corresponds to decline of 53% relative to the second period. Finally, the last period constitutes the months after October 2015. During this period, the average number of complaints is 96 per month, which corresponds to a 43% decline relative to the previous period.

The dashed lines represent the three events we identify as varying the level and type of monitoring. During our period of interest, public perception and focus on policing changed significantly. To contextualize these events, we include frequencies of topics and phrases used in articles between 2007 and 2018 from the *Chicago Tribune*, the city's largest newspaper, relating to the Chicago Police Department¹⁰. The articles were coded to mentions of relevant topics (left figure) and monitoring agencies (right figure)¹¹ in

⁹While weekly data is available, the events cannot be isolated to singular days for two reasons: information transmission about the events may take more than a week, and the events themselves involved smaller and related events around the month of the main event.

¹⁰The initial sample of *Chicago Tribune* articles was the first 8,000 articles available after performing a search on chicagotribune.com/search/ for "Chicago Police Department".

¹¹i.e. agencies that impact the level of oversight, namely police union and the US department of Justice

Figure 2.

The first event, the FOP Note, occurred in November 2009. The second vice-president of the Fraternal Order of Police (FOP)—a member of the Lodge’s Board of Directors—published an article in the union’s monthly newsletter¹² warning officers of the severity and potential risks of being named a witness officer in a complaint investigation (See Figure A.1 in section B). Furthermore, this article appeared right after an unexpected and unfavorable ruling in *Bond v. Chicago*, occurring in early November 2009, which made clear to the union that officer complaint records would be the focus of further lawsuits.

This period is ideal as it occurred before the current debate on police accountability and misconduct. Directly following the *Bond v. Chicago* ruling and the note, there was no change in media coverage of the police department. As illustrated in Figure 2, most of the articles involving the CPD during this period were related to violence in the city, while topics related to policing accountability, misconduct, the DOJ, or the police union were negligible and only appeared with any frequency years later.

The second event, the Kalven Note, was in response to the ruling in *Kalven v. Chicago* (2014). The FOP President published a front-page report in the August 2014 edition of the monthly newsletter informing officers that their complaint records would immediately become available to the public (See Figure A.2 in section B). Again, we see no change in *Chicago Tribune* coverage of the CPD with reference to police accountability, the DOJ, or the police union. However, there is a transitory and relatively small increase in coverage of police misconduct—this may be due to the well-publicized police shooting in Ferguson, Missouri in the same month, which generated a significant amount of nationwide scrutiny of police behavior¹³. If the event in Ferguson impacted Chicago, any effect of the Kalven Note would be attenuated. Since the union newsletter is local and the notes are only relevant to Chicago police officers, spillovers to other police departments in the country are unlikely.

¹²The monthly newsletter contains various announcements, articles of interest for officers, and memos from the union to officers on topics including police related current events, pensions, union bargaining, funerals, and court cases.

¹³The Newsletter from September 2014 contains an article written by an FOP member where he shared his opinion on the event in Ferguson and its implication on policing, political leadership, and media coverage. Aside from this article, we were not able to find any official statement discussing the event in Ferguson in the FOP newsletter.

The third event, the OIS Scandal, is the public release of the dash-cam footage of the killing of teenager Laquan McDonald by a CPD officer. The footage, released in November 2015¹⁴, immediately brought national attention (Davey and Smith [2015]) and scrutiny onto the department (See Figure A.3 in section B). The scandal ultimately triggered a civil rights investigation by the Department of Justice (DOJ) and resulted in the implementation of multiple reforms to the Chicago Police Department and its oversight system by consent decree¹⁵. Figure 2 indicates that the scandal coincides with a spike in the number of articles related to the CPD. Topics related to violence, police accountability, and police misconduct all substantially increased during the month of the shooting leading to a level-increase in all topics in the following months. Moreover, the number of articles mentioning the DOJ or the police union also spiked during this month.

The difference in public scrutiny surrounding these events is significant. Following both the FOP Note and the Kalven Note, news stories related to police accountability and/or the notes themselves were not significant. As such, we believe these events can only induce an increase in self-monitoring among police officers. In contrast, public attention toward police accountability and misconduct spiked following the OIS Scandal. Police self-monitoring increased (likely to avoid further scandal), but there was a simultaneous increase in public-monitoring triggered by the salient example of police misconduct presented in the OIS itself, as well as in the subsequent DOJ Investigation. We exploit this significant difference in order to isolate the effect of oversight.

3 Chicago level analysis

We start our analysis by exploiting micro-level data from Chicago to study the impact of the different types of oversight on crime, officer conduct, and officer effort. The empirical strategy relies on a time series RD (Auffhammer and Kellogg [2011], Anderson [2014]) and captures the effect of oversight in a window around each event of interest.

¹⁴The shooting occurred in October 2014 but drew no media attention until the footage was released.

¹⁵Donohue [2017] provides a detailed discussion of this incident and its implication on policing.

3.1 Data and sample selection

Complaints data

The dataset of misconduct complaints was obtained via Freedom of Information Act (FOIA) requests as part of *Kalven v. the City of Chicago & the Chicago Police Department* (*Kalven v. Chicago*). This data contains the universe of misconduct allegations filed against CPD officers. Each complaint corresponds to a unique complaint registry number, which represents a single incident involving any number of officers. Complainants can initiate their complaint by phone, in person at the oversight agency’s location, by mail, with any CPD supervisor at any district station, or over the internet. Moreover, officers can file a complaint against other officers (See Ba [2018] for more details).

One of the main advantages of the data obtained through *Kalven v. Chicago* is that complaints are categorized by the type of alleged misconduct. We categorize misconduct complaints into four categories (in order of severity): Serious (constitutional violations); Failure to Provide Service (FPS); Other; and Unknown. As a single incident may involve multiple individual allegations, a complaint is identified by its highest level of severity, whether or not any officer was identified in the complaint, and whether or not there was a verdict (finding) on the allegation. Serious complaints are composed of allegations against a known officer involving improper use of force, search, arrest or lock up, or verbal abuse. Allegations in this category tend to be filed by an individual suspected by the police of criminal activity. FPS complaints against a known officer involve a failure to provide service or conduct unbecoming of an officer. FPS allegations tend to be filed by potential victims of a crime or people seeking help from the police¹⁶. “Other” complaints are allegations with a non-missing finding and a known officer where the complaint does not fall into either the Serious or FPS category. “Other” categories of misconduct may include misfiling reports, owing money to the city, operational personnel violations, or workplace complaints. For “Unknown” complaints, all allegations associated with the complaint either have missing findings or no identified officers.

¹⁶Ba [2018] shows civilian complainants have different incentives depending on the category of alleged misconduct.

Crime data

The crime data reflects reported incidents that occurred in Chicago from 2001 to the present. Each incident contains information about the crime, such as location, date, type, and whether or not an arrest was made. For each police beat, we compute the monthly number of total incidents and arrests for index crimes (violent and property crimes) and other crimes (i.e. less serious offenses).

Sample selection

We merge complaints and crime data with demographic information from the 2010 Census. We aggregate the outcomes of interest by beat and month. All the events in the analysis are geocoded according to map [A.4](#) to make events spatially comparable across time, as some police beats changed during 2012. We do not include beats that are located outside of Chicago. Beats that do not have any residents, according to the 2010 Census data, are removed from the sample. The resulting sample contains a total of 276 beats. For each event, we have 34 pre-intervention periods and six post-intervention periods. In other words, the maximum number of periods such that each event can be observed with the same number of pre- and post- periods is 40 months.

As suggested in [Hausman and Rapson \[2018\]](#), we account for seasonality in the outcomes of interests. In other word, the variable y_{bt} corresponds to the residuals that come from regressing the outcome on month dummies and quadratic functions of temperature and precipitation as well as beat fixed-effects.

3.2 Empirical strategy

We estimate how the events impacted complaints, crimes, and arrests using a regression discontinuity design where time is the running variable ([Auffhammer and Kellogg \[2011\]](#), [Anderson \[2014\]](#)). Specifically, for each outcome y in beat b during month-year t , we estimate the following local linear regression

$$y_{bt} = \beta \cdot \text{Monitoring}_{bt} + f(\text{date}_{bt}) + \alpha_b + \varepsilon_{bt} \quad (1)$$

where y_{bt} is the outcome per 10,000 residents. Because of the presence of zeros in the data, we chose to analyze per 10,000 capita levels, rather than percentage changes or logs. The beat fixed effects are given α_b . The variable of interest, Monitoring_{bt} , equals one during or after the month of the event, and zero otherwise. Depending on the specification of interest, Monitoring_{bt} corresponds to the FOP Note, Kalven Note, or OIS Scandal. The function $f(\text{date}_{bt})$ is a flexible function of the date as the number of months from the event of interest (centered at zero). As recommended by [Gelman and Imbens \[2019\]](#) we use local linear and quadratic polynomials function for $f(\text{date}_{bt})$.¹⁷

Identification of our parameter of interest, β , comes from assuming that after controlling for $f(\text{date}_{bt})$ and α_b , the error term, ε_{bt} , is exogenous. In other words, the change in the outcomes was only caused by the discontinuous change in the level of monitoring, i.e. β at $\text{date}_{bt} = 0$, after controlling for time trends and beat fixed-effects. The key assumption necessary to attribute any immediate change in the outcomes at the moment of the event of interest is that no unobserved factor which affects the outcome also systematically changed at the same point in time. We specify a uniform kernel ([Imbens and Lemieux \[2008\]](#)).¹⁸ To make all events comparable, the bandwidth is chosen based on data availability and to maintain the same number of pre- and post-months around the event of interest, which is 34 and 6 months of data in the pre- and post-periods.

The error term ε_{bt} is assumed to have non-constant variance, and standard errors are clustered at the beat level. This assumption allows to account for correlation over time and across beats ([Cameron et al. \[2011\]](#)). We follow [Calonico et al. \[2014, 2017\]](#) to obtain our local polynomial estimators and standard errors.

¹⁷The authors argue that higher-order polynomials in RD design leads to imprecise estimates, poor coverage of confidence intervals, and sensitivity to the degree of the polynomial.

¹⁸We find qualitatively similar estimates using a triangular kernel

3.3 Results

Tables 1, 2, and 3 present the responses of complaints, crimes, and arrests to monitoring, respectively. In each table, Panel A, B, and C report the impact of the FOP Note, Kalven Note, and the OIS Scandal on the outcome of interest, respectively. Odd-numbered columns consider a local linear polynomial function for $f(date_{bt})$, while even-numbered columns consider a quadratic polynomial function. As a companion to the tables, we also provide graphical evidence of our results in Figures 3-11. These figures indicate that the quadratic polynomial function fits the data better.

FOP and Kalven Notes

Panels A and B of Tables 1-3 capture how an increase in self-monitoring impacted police conduct, the crime level, and police effort. Overall, Table 1 indicates that the number of complaints decreases after the notes. Columns (1) and (2) present the results for total complaint rates. The linear specification suggests total complaints dropped by 0.022-0.027 per 10,000 ($p < 0.05$) after the release of the notes (mean at baseline is 0.26 for the FOP Note and 0.20 for the Kalven Note). However, under the quadratic specification, the number of total complaints were not significantly impacted by the Kalven and FOP Notes.

Figures 3 and 4 confirm the abrupt decrease, but the point estimates are within the 90% confidence interval. However, given the wide-ranging forms of misconduct and affected populations types (suspects, victims of crime, or other officers, etc.), pooling all complaints is not informative of changes in policing behavior. In fact, grouping disparate complaint categories may obscure the impact of such events on specific policing behaviors.

Columns (3) and (4) of Table 1 and Figures 3 and 4 indicate that Serious complaints significantly decrease in response to increased self-monitoring. After the FOP Note, the number of Serious complaints by beat dropped by 0.010-0.022 per 10,000 residents ($p < 0.01$), which represents a decrease relative to the baseline mean of at least 11%. After the Kalven Note, Serious allegations per beat dropped by 0.015 to 0.022 per 10,000 residents ($p < 0.01$), which represents a 35% decrease relative to the baseline mean in our preferred specification.

Panel A of Table 1 indicates that the local linear regression suggests that the FOP Note did not impact FPS complaints, while the number of complaints increased by 0.009 per 10,000 residents ($p < 0.1$) under the quadratic specification. The number of FPS allegations declined after the Kalven Note, but the results are not statistically significant.

Columns (7)-(10) of Table 1 show that the notes did not have any significant impact on Other complaints. The null result for Other complaints supports the hypothesis that these events impacted officer conduct with civilians and did not impact the internal workings of the CPD or officers' workplace conduct. Figure 3 and 4 confirm there is no discontinuous change in the number of Other complaints around the time of the event.

Panels A and B of Table 2 present the impact of the notes on crime rates by category: all, violent, property, and other types of crime. The RD results for the notes produced economically small and not statistically significant increases in all types of crime (all increases are less than 5.4% relative to baseline). Figures 6 and 7 confirm that there is no discontinuous change in the number of crimes before and after the release of the notes. These results suggest that the increase in self-monitoring, and the subsequent decrease in serious complaints, did not yield an increase in crime.

Finally, Panels A and B of Table 3 show the impact of the notes on arrests, our proxy for officers' effort. The arrest types correspond to the crime types: all, violent, property, and other. The FOP Note appears to increase all and other arrests ($p < 0.01$ and $p < 0.01$, respectively) under the linear specification, but under the quadratic specification it has a small and insignificant negative effect. Similarly, under the FOP Note, the linear specification finds an increase in violent arrests ($p < 0.01$), while the quadratic specification shows a positive but insignificant effect. Property arrests are not significantly impacted by the FOP Note, though both specifications show modest positive effects. Though these results are not robust, they certainly provide evidence against the proposition that an increase in self-monitoring leads to a decrease in police effort as measured by arrests.

The Kalven Note, by comparison, caused no significant changes to any arrest type. All specifications and crime types produced insignificant, small, and negative results, with the exception of property crime under the quadratic specification with a minuscule positive coefficient (see Fig. 10). These relatively precise and consistent null effects indicate that arrest patterns were not impacted by the release of the Kalven Note, despite a notable

decrease in Serious complaints.

OIS Scandal

We compare the results above with the impact of the OIS Scandal on misconduct, crime levels, and officer effort. Panel C of Table 1 and Figure 8 indicate that, after the OIS Scandal, allegations of misconduct discontinuously increased for all types of complaints, except Other complaints. The total number of complaints sharply increased by about 0.054 per 10,000 residents ($p < 0.01$), which represents a 36.7% increase relative to the baseline mean.

Columns (3) and (4) suggest that the OIS Scandal increased the number of Serious complaints by 50%, with an increase of 0.027 per 10,000 residents ($p < 0.01$). This is the result of two countervailing effects: officers likely decreased their propensity to engage in misconduct with suspects to avoid another scandal, and suspects were more likely to report abuse by police officers. The next two columns of Table 1 suggest that FPS complaint rates increased by 0.014 ($p < 0.01$) from 0.018 at baseline. As with Serious complaints, the OIS Scandal caused a large increase in FPS complaints, indicating that either officer response times and conduct were poorer and/or civilians were less tolerant of such conduct due to the highly public scandal.

The OIS Scandal significantly increased the number of allegations of police misconduct from both suspects and victims of crime, indicating a change in both officer and civilian behavior. Given the simultaneity bias, the multiple mechanisms driving these complaint changes are indistinguishable. For example, civilians may be more willing to come forward when law enforcement is scrutinized, and officers may be less likely to police aggressively; these mutually confounding mechanisms on complaint rates would have contradicting effects but their occurrence would be simultaneous.

Panel C of Table 2 indicates that crime rates increase by 0.64-0.95 per 10,000 residents ($p < 0.05$) after the OIS Scandal, relative to a baseline mean of 11.85. The local linear specification indicates that the number of violent crimes increases by 12.3% relative to the baseline mean (0.123 per 10,000 residents), while the coefficient of interest is smaller and no longer significant when using the quadratic specification (0.123 vs. 0.044 per 10,000

residents).

Columns (5) and (6) of Table 2 suggest that property crime drove most of this increase (note that non-violent crimes represent about 92% of the offenses); none of the coefficients in these columns were significant for either the FOP or Kalven Notes. After the OIS Scandal, the linear and quadratic specifications indicate that there was a 15% and 7% increase in property crime ($p < 0.01$ and $p < 0.05$), respectively. Columns (7) and (8) indicate that the OIS Scandal increased the number of other crimes, but the results are only significant at the 10% level for the quadratic specification.

Panel C of Table 2 indicates that under the OIS Scandal almost all arrest types and specifications produced insignificant and mostly negative results. The exceptions being all arrests increasing by 8.8% ($p < 0.1$) and Other arrests increasing by 10% ($p < 0.05$), both under the quadratic specification.

In comparison to the impacts on complaints and crime, none of the events had conclusive impacts on the level of any type of arrest. Under the first two events, where crime did not change significantly, this would indicate that policing was not negatively impacted, while serious misconduct declined significantly in response to more self-monitoring. Under the OIS Scandal, on the other hand, the increase in misconduct allegations and increased level of crime coupled with no corresponding increase in arrests indicates a reduction in police effort. However, this result cannot be interpreted as a response to increased self-monitoring because the scandal simultaneously changed civilian incentives as indicated by the increase in complaints.

4 Chicago vs Rest of the U.S.

We complete our analysis by looking at the impact of oversight on crime outcomes by comparing Chicago to other major cities in the U.S.. The panel structure of the data allows us to control for Chicago's unobserved characteristics and common trends that it shares with other major cities. We attempt to control for these factors to avoid omitted variable bias.

4.1 Data

4.1.1 Sources and sample selection

The primary data source for this analysis is the UCR data for offenses known and clearances by arrest from 1960 to 2016. This dataset has been collected by [Kaplan \[2018\]](#), who constructed a monthly dataset on a police department-level basis for the following index crimes: murder, forcible rape, aggravated assault, battery, robbery, burglary, larceny, and motor vehicle theft. We supplement those datasets with monthly average temperature and precipitation measurements for each county using data from the Daily Global Historical Climatology Network (GHCN) ([Menne and Houston \[2012\]](#)).

Our analysis of crime data focuses on the period between January 2007 and December 2016. We focus on the Chicago Police Department and local police departments from cities with populations exceeding 250,000 at any time between 2007 and 2016 in the UCR data. Since Chicago is the third-largest city in the U.S., our preferred specification focuses on cities with populations larger than 500,000, as this provides a more appropriate donor pool to construct a counterfactual Chicago.

The main outcome of interest is the number of reported crimes per 100,000 residents (i.e. crime rate) in a municipality. Because of the presence of zeros in the data, especially for murder, we chose to analyze per 100,000 capita levels rather than logs or percentage changes. In order to account for seasonal patterns in crime outcomes, we consider the residualized crime rates by regressing crime rates on month dummies and using the residuals as the dependent variable.

For each department in our sample, we construct reported monthly crime rates for violent crimes (murder, robbery, aggravated assault) and property crimes (burglary, larceny, motor vehicle theft). Because the UCR does not systematically provide the number of reported rapes for Chicago, we do not include this outcome in our analysis.

The number of arrests is an alternative outcome of interest that has been considered by economists ([Mas \[2006\]](#), [Shi \[2009\]](#)). This measure can potentially capture the level of police effort or activity to minimize the level of crimes. Unfortunately, the UCR data does not systematically provide the number of crimes cleared by arrest for the Chicago Police Department.

As noted in [Maltz and Weiss \[2006\]](#), [Evans and Owens \[2007\]](#), [Chalfin and McCrary \[2018\]](#), [Mello \[2019\]](#), the UCR data contains reporting errors that vary across agency. Hence, the data requires thorough cleaning before use. We follow [Mello \[2019\]](#)’s regression-based approach to identify record errors and extreme outliers (See Appendix [A](#) for more details). We delete agencies that are flagged by our procedure as outliers. Following [Chalfin and McCrary \[2018\]](#) and [Mello \[2019\]](#), we use a smoothed version of the measure for the city population reported in the UCR.

We exclude agencies that report negative crimes and/or police departments that did not report any crimes in a given year. We only consider local police departments for our analysis (excluding sheriff’s departments, U.S. Marshals, state police, and other special jurisdictions). Agencies with missing population data are also dropped.

Keeping cities that have been subject to major instances of public scrutiny and monitoring in our control group may impact our results. For instance, the DOJ opened an investigation into the Baltimore Police Department (BPD) in May 2015. The DOJ evaluated whether or not BPD engaged in a pattern or practice of conduct that violated the US Constitution or federal law. Our preferred specification restricts the donor pool to police departments that have never been investigated by the DOJ under [the Section 14141 of the Violent Crime Act of 1994](#) (“pattern-or-practice cases” or “14141 cases”) in the control group (as well as requiring populations of at least 500,000). We discuss the impact of including cities that were subject to DOJ investigation in Section [5](#).¹⁹

The resulting data set contains a panel of the Chicago Police Department and 24 to 67 additional local police agencies from January 2007 to December 2016, depending on the restrictions. As discussed previously, our main specification only includes large jurisdictions (population $\geq 500,000$) that were not subject to DOJ investigation. Table [A.1](#) and Appendix [A](#) explain the sample selection criteria in more detail.

¹⁹During the time period of interest, the DOJ was particularly active and interested in police misconduct cases. As a result, most jurisdictions subject to heavy public monitoring were simultaneously subject to DOJ investigation.

4.1.2 Summary Statistics

Table 4 contains the summary information of Chicago compared to the rest of the US in the sample from 2007 through 2016. Chicago’s population is more than two and a half times that of the average city in the control group. Time-varying controls are present at the county and monthly level such as average temperature and average precipitation. Cities in the control group tend to be warmer and to have less precipitation than Chicago.

Crime rates (per 100,000 residents) are observed at the department level and aggregated by month. Violent crime is comprised of aggravated assault and battery (assault), murder, and robbery. While Chicago has an average assault rate less than one-third of the rest of the U.S., Chicago’s murder and robbery rates are almost double those of the rest of the country. Violent crime rates in the rest of the U.S. are double those in Chicago, primarily due to the higher levels of assault in other jurisdictions. The rate of property crimes is only about 5% lower in Chicago than elsewhere, and rates of burglary and larceny tend to be higher in the rest of the U.S., while rates of motor vehicle theft tends to be higher in Chicago. Figures 12 plot the trends in violent and property crime rates in Chicago and the rest of the United States. These figures suggest that the rest of the United States may not provide a suitable comparison group for Chicago to study the effects of the three events of interest on crimes.

Given the drastically different composition of violent crime in other donor cities, the synthetic control results for violent crime will overweigh assault in the construction of the predicted path for the synthetic Chicago. As a result, we rely mostly on the more detailed categories of violent crime for reasonable results.

4.2 Empirical Strategy

4.2.1 Setup

Our main goal is to understand how the increases in monitoring, identified by the three events which induced variation in misconduct allegations, impacted crime rates in Chicago. Our main challenge is to construct three counterfactuals for the city around the time of each event. Let I_t be a post-treatment dummy and t_0 be the date of treatment. Using the

potential outcomes framework for causal inference (Neyman [1923], Rubin [1974], Holland [1986]), let $y_{ct}(0)$ and $y_{ct}(1)$ be the potential outcomes for agency c at time t when $I_t = 1(t < t_0) = 0$ and $I_t = 1(t \geq t_0) = 1$, where treatment indicates occurrence of one of these events. The potential outcomes are given by $y_{ct}(0) = x'_{ct}\gamma + f'_t\lambda_c + \varepsilon_{ct}$ and $y_{ct}(1) = \beta_{ct} + x'_{ct}\gamma + f'_t\lambda_c + \varepsilon_{ct}$. The vector x_{ct} contains time varying controls. Time-varying and agency specific unobservables are given by f_t and λ_c , respectively.

The agency treatment effect at time t is $\beta_{ct} = y_{ct}(1) - y_{ct}(0)$ for any agency c such that $t \geq t_0$. Because $y_{ct}(1)$ is observed for treated units in post-treatment periods, the main objective is to construct counterfactuals for each treated unit in post-treatment periods, i.e., $y_{ct}(0)$ for treated c and $t \geq t_0$. Potential outcomes and the value of the treatment determine the observed outcome such that

$$y_{ct} = y_{ct}(0) \cdot (1 - D_c I_t) + y_{ct}(1) \cdot D_c I_t$$

where D_c equals one if the agency is Chicago, and zero otherwise. We use a combination of other U.S. police departments from the 24 to 67 municipalities discussed above to construct a counterfactual control police department which resembles the Chicago Police Department with respect to the relevant crime characteristics before the three events of interest. We construct a panel of 48 months for each event to compare the different treatment effects. The number of months, $T = 48$, corresponds to the maximum number of periods such that each event can be observed with the same number of pre- and post-periods. In each panel, date $t_0 = 0$ corresponds to the date of the event, and each panel has 34 pre-intervention (T_0) periods and 14 post-intervention periods (T^*).

4.2.2 Estimation

The treatment effect for the event of interest could be recovered using a difference-in-difference approach, however visual evidence from Figure 12 suggests the “parallel trends” assumption appears to be violated for some of our outcomes. Another approach to build a counterfactual group for Chicago would use the synthetic control method from Abadie and Gardeazabal [2003], Abadie et al. [2010, 2015]. This method constructs a weighted average of control units that matches the treated unit’s pre-treatment outcomes. How-

ever, the standard synthetic control method does not allow for multiple treated units and treatment periods. We adopt the interactive fixed effects model from Xu [2017], which generalizes the synthetic control method to cases of multiple treated units and treatment periods. This model builds on the interactive fixed effect models from Bai [2009]. Athey et al. [2018] generalize interactive fixed effects model (Bai [2009], Gobillon and Magnac [2016], Xu [2017]), allowing for matrix completion. Following Xu [2017], treated and untreated units of the outcome variable are given by:

$$y_{ct} = \beta_{ct} D_c I_t + x'_{ct} \gamma + f'_t \lambda_c + \varepsilon_{ct} \quad (2)$$

where f_t are time-varying unobserved factors, λ_c are agency-specific factor loadings. The number of factors is L . The variable D_c equals one if the agency is Chicago, and zero otherwise; and $I_t = 1(t \geq t_0)$ is a time indicator of treatment, where $t_0 = 0$ is the date of the event. This model nests the traditional linear difference-in-differences model, which is obtained when $\lambda_c = (\alpha_c, 1)'$ and $f_t = (1, \tau_t)'$, such that $f'_t \lambda_c = \alpha_c + \tau_t$. Following Bai [2009], we assume that ε_{ct} is uncorrelated with (x_{ct}, f_t, λ_c) , as well as other identification conditions in Bai [2009] and Gobillon and Magnac [2016].

In order to construct a counterfactual group for Chicago, Xu [2017] proposed to first estimate equation 2 using the control group data to recover $(\hat{\gamma}, \hat{f}_t, \hat{\lambda}_c)$. Second, using $(\hat{\gamma}, \hat{f}_t)$, recover the factor loadings for each treated unit ($D_c = 1$) by minimizing the mean squared error of the predicted treated outcome in pre-treatment periods ($t < t_0$), i.e.:

$$\lambda_c = \underset{\lambda_c}{\operatorname{argmin}} \sum_c \sum_t (y_{ct}^0 - x_{ct}^{0'} \hat{\gamma} - \hat{f}_t^{0'} \lambda_c) (y_{ct}^0 - x_{ct}^{0'} \hat{\gamma} - \hat{f}_t^{0'} \lambda_c)'$$

where y_{ct}^0 , x_{ct}^0 , and f_t^0 correspond to the outcome, controls, and time-varying factors during the pre-intervention periods, i.e. $t < t_0$. Finally, construct counterfactual groups using $(\hat{\gamma}, \hat{F}, \hat{\lambda}_c)$ and the treatment effect, $\hat{\beta}_t$, such that

$$\begin{aligned}
\text{Counterfactual : } \hat{y}_{ct}(0) &= x'_{ct}\hat{\gamma} + \hat{f}'_t\hat{\lambda}_c, & \text{for } D_c = 1 \text{ and } t \geq t_0 \\
\hat{\beta}_{ct} &= y_{ct}(1) - \hat{y}_{ct}(0), & \text{for } t \geq t_0 \\
\hat{\beta}_t &= \frac{1}{N_T} \sum_{c=1}^{N_T} \hat{\beta}_{ct}, & \text{for } t \geq t_0
\end{aligned}$$

where N_T is the number of treated groups. In practice the exact number of factors L is unknown. [Xu \[2017\]](#) develops a cross-validation procedure to determine the exact number of factors. The proposed method relies on control group information and information from the treatment group before the intervention.

Standard errors for the generalized synthetic controls methods (GSC) are based on parametric bootstraps (blocked at the agency level) of 10,000 replications. [Xu \[2017\]](#) provides details about the procedure but leaves a formal justification for future research. The proposed method obtains uncertainty estimates conditional on observables, unobserved factors, and factor loadings using a parametric bootstrap procedure by resampling the residuals. Resampling the residuals preserves the serial correlation within the units ([Beck and Katz \[1995\]](#), [Xu \[2017\]](#)). Details about the procedure can be found in [Xu \[2017\]](#). Although there is no formal justification for this procedure ([Chernozhukov et al. \[2017\]](#)), we implement this method to conduct inference on our estimates.

4.3 Results

This section presents our main GSC results. As discussed previously, our main analysis only considers cities with a populations of at least 500,000 which have not been subject to DOJ investigation. Table 5 summarizes the effects of oversight on crime rates for the three events when using the generalized synthetic control methods. For each outcome, we report the mean of the dependent variable in Chicago during the pre-intervention period. We account for seasonal patterns in crime outcomes by removing the month effect and report the residuals of this regression.

FOP Note

We present the results in Column (1) of Table 5. The FOP Note decreased violent crimes while increasing property crimes. After this event, the number of violent crimes decrease by 12.53 per 100,000 residents ($p < 0.05$), which represents a drop of 12.7% relative to the baseline mean in Chicago. The number of assaults and robberies dropped by 5.4 and 2.9 per 100,000 residents, while the murder rates increased by 0.057; however, none of the effects on the different violent crime categories are statistically different from zero. The increase in property crime seems to be driven by a significant increase in motor vehicle theft of 8.6 per 100,000 residents ($p < 0.1$), i.e., a 16% increase compared to the baseline period in Chicago. Other property crimes, such as larceny and burglary, do not seem to be impacted by the event. While robbery and assault decrease, murder is the only violent crime that increases after the release of the note.

The pre-trends in crimes per 100,000 residents, presented in Figure 13, indicate that both Chicago and its counterfactual exhibit similar patterns. Most of the outcomes seems to have similar trends after the release of the FOP Note. The number of violent crimes in Chicago decreases relative to its counterfactual, while the number of motor vehicle thefts increases after the union sent the note.

Figure 14 presents the gap in trends between Chicago and its counterfactual before and after the FOP Note. The number of violent crimes declines after the release of the FOP Note, though this finding is not statistically significant. The results by type of violent crime separately remain close to zero and statistically non-significant. The FOP Note increased the number of motor vehicle theft, but the other property crimes (larceny and burglary) were not impacted.

Kalven Note

With the exception of the murder rate, Column (2) of Table 5 suggests that rates of all types of crime decreased after the Kalven Note. However, robbery is the only variable that was significantly impacted by the note ($p < 0.05$). The reduction in robberies represent a drop of 4.9 per 100,000 residents (mean at baseline is 37.4)

Figure 15 plots the trends in crimes per 100,000 residents for Chicago and its counter-

factual before and after the Kalven Note. Chicago and its counterfactual display similar trends during the pre-intervention period for all outcomes. The figures show high variance in Chicago's murder time-series in both the pre- and post-intervention periods. In the months following the Kalven Note, the number of motor vehicle thefts increased in the counterfactual groups, whereas it remained stable in Chicago. The number of robberies displays a similar pattern. Compared to Chicago, the number of assaults increases in the counterfactual groups in the five months following the publication of the Kalven Note.

Likewise, the total number of violent crimes increases in the counterfactual groups, a result that is likely driven by the aforementioned increase in assault and the overweighting of assault for the synthetic Chicago with respect to aggregated violent crimes. The trends in violent crimes and assault between the two groups converge six months after the Kalven Note. As previously mentioned, because the Kalven Note was published in the same month as the nationally-covered event in Ferguson, it is plausible that cities in the counterfactual group were more impacted by the events in Missouri than was Chicago.

Figure 16 presents the gap in trends between Chicago and its counterfactual before and after the Kalven Note. The gap between the number of property crimes and incidents of larceny, burglary, and murder are relatively flat and close to zero before and after the event, as is the gap between the number of violent crimes and assaults. A negative spike in violent crimes and assault in the month after the event is not significant. Motor vehicle thefts and robberies seem to decline after the Kalven Note, but the confidence intervals are too wide to reject the null hypothesis.

OIS Scandal

Crime rates in Chicago increase after the OIS Scandal. Except for motor vehicle theft, the OIS Scandal tends to have a larger effect (in absolute value) on crime than either the FOP or the Kalven Note. Column (3) of Table 5 indicates that violent crimes increased by about 8 per 100,000 residents after the OIS Scandal. Assault increased by 11 per 100,000 residents, which corresponds to a 29% increase compared to the baseline period in Chicago, but this increase is not statistically significant. The number of robberies increased by

5.83 (mean at baseline is 31.8) per 100,000 residents, and the effect is significant at the 10% level. The number of murders significantly increased by 0.68 per 100,000 residents ($p < 0.01$). In other words, relative to the counterfactual Chicago, there were 18.4 more murders each month in the period post-OIS Scandal, as opposed to 1.4-1.6 ($p > 0.1$) for the FOP and Kalven Notes. The number of property crimes increased by 28 per 100,000 residents ($p < 0.05$). The scandal had a positive effect on burglary and larceny crimes and a small negative effect on the number of motor vehicle thefts, but none of these effects were statistically significant.

Figure 17 plots the trends in crimes per 100,000 residents for Chicago and its counterfactual before and after the OIS Scandal. For most of the outcomes, Chicago and its counterfactual display similar trends during the pre-intervention period. The only exception is for motor vehicle theft, which has different trends between Chicago and its counterfactual during months -34 and -32. The post-intervention period suggests that both the number of violent and property crimes increased. These figures confirm that murder and robbery rates differ after the OIS Scandal.

Figure 18 presents the gap in trends between Chicago and its counterfactual before and after the OIS Scandal. We find that the number of violent crimes per 100,000 residents increased after the scandal, but that the effect has a relatively wide confidence interval and is not statistically significant. There is a statistically significant increase in the number of murders and robberies for some periods post-event. The results for property crimes are too imprecise to draw any conclusions.

5 Robustness

5.1 Chicago level analysis

The validity of our RD design requires that the conditional expectation $E[\varepsilon_{bt} | date_{bt}]$ is smooth as $date_{bt}$ crosses the RD threshold (Anderson [2014]). One threat to our design could be that officers or civilians alter their behavior in anticipation to a change in monitoring. Although, the exact timing of the three events are plausibly exogenous, it is important to rule out any possibility of anticipation or avoidance effects.

Tables A.3-A.11 estimate a “donut” RD by removing observations near the threshold (Barreca et al. [2011]) to evaluate if our estimates are sensitive to observations close to the cutoff. We consider three donut RDs and remove: (1) the month of the event, (2) the month of the event and the preceding month, and (3) the month of the event and two preceding months. We find that the results remain qualitatively similar to those found in Section 3.3.²⁰

We run a series of placebo tests. Each placebo estimate first assigns a window around a false event date, and then uses an RD to estimate the effect of oversight on complaints, crimes, and arrests. To avoid having the placebo estimates be influenced by any jump at the true cutoff, the placebo window is between from April 2013 to January 2014. The window is chosen so that: (1) it does not include the six months post-event, and (2) we have 34 months preceding the placebo event, and six months following the placebo date. Note that our placebos are still partially treated by the FOP Note as we do not have enough data before that event.

Figures A.5-A.7 graph the true and placebo RD estimates. We also report the 10th and 90th percentiles of our preferred specification. As the graphs make clear, we see that the number of Serious complaints significantly decreased after the release of FOP Note, and increased after the OIS Scandal. FPS complaints also increased significantly after the scandal. These results support the main conclusions from our original RD.

5.2 Chicago vs Rest of the U.S

Alternative control groups As discussed previously, our primary analysis focuses on cities that have not been subject to DOJ investigation and with populations of at least 500,000. Table A.12 provides the results of the supplemental analysis.

We consider three additional sample constructions for the donor pool. Sample 1 corresponds to cities with at least 250,000 people. Sample 2 corresponds to cities with at least 250,000 people but excludes cities that were subject to DOJ investigation at some point

²⁰Our choice of bandwidth is limited since we only have 5 months of complaints data after the OIS Scandal, therefore we cannot perform bandwidth selection adjustments for few mass points in the running variable in the spirit of Imbens and Kalyanaraman [2012].

since 1994.²¹ Sample 3 is cities with a population of at least 500,000. Both Sample 1 and Sample 3 permit cities with prior DOJ investigations.

Overall, results from Table A.12 are qualitatively similar to those drawn from Table 5. We do not find evidence that self-monitoring increases crime, while the OIS Scandal resulted in a significant increase in murders and robberies. The alternative sample selections indicate that including cities which were investigated by the DOJ and those with populations smaller than 500,000 residents produced less precise estimates from our preferred sample. The GSC estimation seems to be sensitive to the composition of the donor pool sample.

Synthetic Control Methods As a complementary approach to the general synthetic control methods, we turn to the synthetic control methods developed in Abadie and Gardeazabal [2003], Abadie et al. [2010, 2015]. The main idea of this method is to construct a synthetic match for Chicago by using the agencies in the control group (i.e. the rest of the U.S.) in such a way that the synthetic control has similar behavior to Chicago before the event of interest. The synthetic control (counterfactual Chicago) is constructed as a weighted average of control units in order to match pre-intervention outcomes for the treated units. The weights are restricted to be non-negative and sum to one. Inference is based on placebo permutation to test if the results are driven by chance using Jones and Marinescu [2018]. Table A.13 summarizes the effects of oversight on crime rates for the three type events when using synthetic control methods.

Overall, these results are qualitatively similar to those drawn from Table 5. Most of the results are statistically nonsignificant when using synthetic control methods and inference using placebo permutations.

The FOP Note increases the number of motor vehicle thefts by about 10 per 100,000 residents ($p < 0.05$). With the exception of murder, the Kalven Note is associated with decreased crime rates.

This table also suggests that the number of crimes went up for all outcomes after the OIS Scandal. The scandal had a significant impact on murder and robbery rates ($p < 0.1$), and the estimated effects are higher than the 95th percentiles of the placebo distribution.

²¹Corresponding to the implementation §14141 of the Violent Crime Act of 1994

The remaining outcomes are within the 5th and 95th percentiles of the placebo distribution.

6 Discussion

We provide evidence which both expands on and reinterprets previous research on the impact of policing the police. Our findings comport with research, such as ([Prendergast \[2001\]](#), [Shi \[2009\]](#), [Heaton \[2010\]](#)), which finds that during periods of major reform following police scandals, when public scrutiny is unusually high and police officers are most cautious, crime rises and police effort declines. We find that following an officer-involved shooting (OIS) scandal, murder and robbery increase but arrests, a proxy for officer effort, do not. These results, along with previous research, would appear to confirm the “Ferguson Effect” theory— that increased scrutiny of police tactics and increased political pressure to prosecute police misconduct leads to reduced effort by officers and causes crime to rise. We provide novel evidence that confirms public scrutiny increases during scandals, by showing that civilian complaints increase significantly after the OIS, despite unchanged arrest levels in response to increased crime.

We interpret these results, in the context of a simultaneity bias inherent in measuring the impact of a scandal as a substitute for oversight on policing outcomes. Simply put, periods of public outrage are not representative of an oversight regime, due to inflated public-monitoring and social unrest. Oversight reforms are largely targeted at internal reform and increased self-imposed monitoring by officers, meaning they do not rely on heightened public vigilance and public attention to reduce misconduct. Furthermore, civilian incentives change during public scandals such that, even in the absence of officers’ increased restraint, potential criminals may be simultaneously responding to general social unrest and mistrust of police by committing more crimes. Distrust in communities hampers crime reporting ([Desmond et al. \[2016\]](#)), which may, holding enforcement constant, increase criminal opportunities. As such, we argue that scandal-based estimates of the effect of oversight on outcomes such as crime and police effort are contaminated by a simultaneity bias due to the simultaneous changes in civilian and officer incentives which

would both, independently, increase crime.

By studying increases in self-monitoring during periods without public scandals, we isolate the impact of oversight on misconduct, crime, and effort. We find that notes from executives of the police union cautioning officers about the seriousness of misconduct allegations, largely unknown to the public, induce a reduction in complaints, particularly allegations of constitutional rights violations, without an increase in crime or a change in arrests. Our analysis demonstrates that police officers are highly responsive to internal oversight mechanisms, even when they do not carry specific incentive shifts or heightened negative consequences. This complements the work of [Mummolo \[2018\]](#), which finds that officers are highly responsive to internal directives and that increased oversight by commanders can improve police productivity.

Moreover, scandal-based studies have thus far focused on short-term outcomes. It is possible that long-term crime rates and rates of misconduct are lower following reforms once public outrage has subsided, due to increased effectiveness, better culture, and more community trust. The long-run effects of public scandals that lead to reform and increased monitoring provide a more nuanced story. Police scandals in the US have triggered DOJ investigations and police reforms under §14141 of the Violent Crime Act of 1994 ([DOJ \[2017\]](#)). Studies have shown that agreements, resulting from major reforms brought about by DOJ investigations, generally were successful in terms of implementing effective and constitutional policing practices and improving police-community relations in the long run. For instance, [Stone et al. \[2009\]](#) found that the Los Angeles Police Department consent decree improved both the quantity and quality of policing while reducing crime. [Chanin \[2015\]](#) studies the effect of police reforms in Pittsburgh, Washington D.C., and Cincinnati, and finds that reforms increased accountability practices and reduced police misconduct without increasing crime.

Given our results, a potential solution for policymakers is to consider proactive policies that will increase police monitoring, without waiting for a public scandal to occur. Such proactive policies should engage key stakeholders; in particular, police unions can act as powerful agents for increasing officers' self-monitoring. Proactive accountability policies may also reduce the likelihood of a major officer-involved-shooting scandal that could lead to civilian distrust, increased crime, and major settlement costs to tax payers

(Schwartz [2014, 2016], Rozema and Schanzenbach [2018]). Self-monitoring may induce better behavior from officers toward civilians, leading to increased legitimacy of legal authorities and heightened community trust. This directly and indirectly increases civilian compliance with the law and cooperation in investigations, as well as making it more difficult for civilians to engage in criminal activities (Skogan and Frydl [2004], Tyler et al. [2015], Desmond et al. [2016]).

Our research has the following limitations. First, we focus primarily on short-run outcomes, as a result of data availability constraints. However, to the best of our knowledge, our work is the first to link the impact of oversight on different types of misconduct, crime, and arrests. Second, the qualitative nature of our events means we cannot directly measure the level of oversight at any given time. Third, both our events are memos from the police union; while they were effective in increasing self-monitoring, other internal reforms, such as introducing more robust oversight regimes, may be differentially impactful on policing outcomes. Nevertheless, we maintain our results show there is a key distinction between policing outcomes under oversight and public outrage.

7 Conclusion

The tradeoff between police oversight and effective policing is documented in a literature largely focusing on periods of high public scrutiny of the police and subsequent reforms (Prendergast [2001], Shi [2009], Heaton [2010]). Such scandals induce officers to increase self-monitoring as a form of self-protection against misconduct accusations and further scandal. However, as a result of the civil unrest and civilian distrust following a scandal, civilians simultaneously become more critical of officer actions and less likely to obey laws and authorities. We provide evidence that estimates of oversight's effect on crime levels during scandals are contaminated by this simultaneity bias.

To avoid this issue, we identify two events that only impacted Chicago police officers' level of self-monitoring by increasing their perceived cost of a misconduct allegation. The first event, an article in the Chicago police union newsletter following an unfavorable court ruling, occurred in 2009, years before the national spotlight turned to police mis-

conduct. The second event, another police union note following the decision in *Kalven v. City of Chicago*, occurred in 2014. Both events induced little, if any, public scrutiny, as evidenced by the absence of a change in reporting topics around policing in the *Chicago Tribune*.

Both of the notes significantly decreased the number of complaints against officers for constitutional violations and had no impact on complaints related to workplace conduct, indicating officers solely increased self-monitoring during interactions with civilians. Furthermore, we find no evidence of increased crime rates across all types of crime as well as no impact on arrests. Such results indicate that the increase in self-monitoring, in the absence of an increase in public-monitoring and its complementary civilian distrust, improved officer conduct while having no adverse impact on public safety.

There are three key takeaways from this paper. First, evaluating the impact of a change in police oversight on policing is challenging as it can impact both civilian and officer behavior. Second, increasing officer self-monitoring through cautionary notes or other “encouragement” by police unions are likely to be effective avenues for decreasing misconduct while maintaining effective policing and public safety. Policymakers should take into consideration the effectiveness of internal directives and the union’s role in influencing officer conduct in the absence of public pressure. Lastly, while this paper offers empirical evidence for proposed police incentives, there is need for a more comprehensive theoretical model of police officers’ objective function as well as that of civilian interactions with officers.

References

- Alberto Abadie and Javier Gardeazabal. The economic costs of conflict: A case study of the basque country. *American economic review*, 93(1):113–132, 2003.
- Alberto Abadie, Alexis Diamond, and Jens Hainmueller. Synthetic control methods for comparative case studies: Estimating the effect of california’s tobacco control program. *Journal of the American statistical Association*, 105(490):493–505, 2010.

- Alberto Abadie, Alexis Diamond, and Jens Hainmueller. Comparative politics and the synthetic control method. *American Journal of Political Science*, 59(2):495–510, 2015.
- Michael L. Anderson. Subways, strikes, and slowdowns: The impacts of public transit on traffic congestion. *American Economic Review*, 104(9):2763–96, September 2014.
- Kate Antonovics and Brian G. Knight. A new look at racial profiling: Evidence from the boston police department. *The Review of Economics and Statistics*, 91(1):163–177, February 2009.
- Shamena Anwar and Hanming Fang. An alternative test of racial prejudice in motor vehicle searches: Theory and evidence. *American Economic Review*, 96(1):127–151, March 2006.
- Susan Athey, Mohsen Bayati, Nikolay Doudchenko, Guido Imbens, and Khashayar Khosravi. Matrix completion methods for causal panel data models. Technical report, National Bureau of Economic Research, 2018.
- Maximilian Auffhammer and Ryan Kellogg. Clearing the air? the effects of gasoline content regulation on air quality. *American Economic Review*, 101(6):2687–2722, October 2011.
- Bocar A Ba. Going the extra mile: the cost of complaint filing, accountability, and law enforcement outcomes in chicago. *Working paper*, 2018.
- Jushan Bai. Panel data models with interactive fixed effects. *Econometrica*, 77(4):1229–1279, 2009.
- Alan I. Barreca, Melanie Guldi, Jason M. Lindo, and Glen R. Waddell. Saving Babies? Revisiting the effect of very low birth weight classification*. *The Quarterly Journal of Economics*, 126(4):2117–2123, 10 2011.
- Nathaniel Beck and Jonathan N. Katz. What to do (and not to do) with time-series cross-section data. *The American Political Science Review*, 89(3):634–647, 1995.

- Gary S Becker and George J Stigler. Law enforcement, malfeasance, and compensation of enforcers. *The Journal of Legal Studies*, pages 1–18, 1974.
- Jean-Pierre Benoit and Juan Dubra. Why do good cops defend bad cops? *International Economic Review*, 45(3):787–809, 2004.
- Renee Binder, Amy Friedli, and Elena Fuentes-Afflick. The new academic environment and faculty misconduct. *Academic Medicine*, (91):175–179, 2016.
- Jordi Blanes i Vidal and Tom Kirchmaier. The effect of police response time on crime clearance rates. *The Review of Economic Studies*, 85(2):855–891, 2017.
- Sebastian Calonico, Matias D. Cattaneo, and Rocio Titiunik. Robust nonparametric confidence intervals for regression-discontinuity designs. *Econometrica*, 82(6):2295–2326, 2014.
- Sebastian Calonico, Matias D. Cattaneo, Max H. Farrell, and Rocio Titiunik. Rdrobust: Software for regression-discontinuity designs. *The Stata Journal*, 17(2):372–404, 2017.
- A. Colin Cameron, Jonah B. Gelbach, and Douglas L. Miller. Robust inference with multiway clustering. *Journal of Business Economic Statistics*, 29(2):238–249, 2011.
- Richard Cassell, Paul G. Fowles. What caused the 2016 chicago homicide spike: an empirical examination of the aclu effect and the role of stop and frisks in preventing gun violence symposium: Federal responses to police misconduct: Possibilities and limits. *University of Illinois Law Review*, 2018:1581, 2018.
- Aaron Chalfin and Justin McCrary. Are us cities underpoliced? theory and evidence. *Review of Economics and Statistics*, 100(1):167–186, 2018.
- Joshua M. Chanin. Examining the sustainability of pattern or practice police misconduct reform. *Police Quarterly*, 18(2):163–192, 2015.
- Cheng Cheng and Wei Long. Improving police services: Evidence from the French Quarter Task Force. *Journal of Public Economics*, 164(C):1–18, 2018.

- Victor Chernozhukov, Kaspar Wuthrich, and Yinchu Zhu. An Exact and Robust Conformal Inference Method for Counterfactual and Synthetic Controls. Papers 1712.09089, arXiv.org, December 2017.
- Monica Davey and Mitch Smith. Video of chicago police shooting a teenager is ordered released. *The New York Times*, 2015.
- Matthew Desmond, Andrew V Papachristos, and David S Kirk. Police violence and citizen crime reporting in the black community. *American Sociological Review*, 81(5):857–876, 2016.
- Dhammika Dharmapala and Thomas J. Miceli. Search, Seizure and (False?) Arrest: An Analysis of Fourth Amendment Remedies when Police can Plant Evidence . *Working Paper*, 2012.
- Rafael Di Tella and Ernesto Schargrodsky. Do police reduce crime? estimates using the allocation of police forces after a terrorist attack. *American Economic Review*, 94(1):115–133, 2004.
- DOJ. The civil rights division’s pattern and practice police reform work: 1994-present. Technical report, Civil Rights Division U.S. Department of Justice, January 2017.
- John J. Donohue. Comey, trump, and the puzzling pattern of crime in 2015 and beyond. *Columbia Law Review*, 117(5):1297–1354, 2017.
- John J. Donohue and Steven D Levitt. The impact of race on policing and arrests. *The Journal of Law and Economics*, 44(2):367–394, 2001.
- Mirko Draca, Stephen Machin, and Robert Witt. Panic on the streets of london: Police, crime, and the july 2005 terror attacks. *American Economic Review*, 101(5):2157–81, 2011.
- Steven N. Durlauf. Assessing racial profiling. *The Economic Journal*, 116(515):F402–F426, 2006.
- William N. Evans and Emily G. Owens. COPS and crime. *Journal of Public Economics*, 91(1-2):181–201, February 2007.

- Chao Fu and Kenneth I Wolpin. Structural Estimation of a Becker-Ehrlich Equilibrium Model of Crime: Allocating Police Across Cities to Reduce Crime. *The Review of Economic Studies*, 85(4):2097–2138, 11 2017.
- Andrew Gelman and Guido Imbens. Why high-order polynomials should not be used in regression discontinuity designs. *Journal of Business & Economic Statistics*, 37(3):447–456, 2019.
- Laurent Gobillon and Thierry Magnac. Regional policy evaluation: Interactive fixed effects and synthetic controls. *Review of Economics and Statistics*, 98(3):535–551, 2016.
- Catherine Hausman and David S. Rapson. Regression discontinuity in time: Considerations for empirical applications. *Annual Review of Resource Economics*, 10(1):533–552, 2018.
- Paul Heaton. Understanding the effects of antiprofiling policies. *The Journal of Law and Economics*, 53(1):29–64, 2010.
- Paul Heaton, Priscillia Hunt, John MacDonald, and Jessica Saunders. The short- and long-run effects of private law enforcement: Evidence from university police. *The Journal of Law and Economics*, 59(4):889–912, 2016.
- Paul W Holland. Statistics and causal inference. *Journal of the American statistical Association*, 81(396):945–960, 1986.
- Jordi Blanes i Vidal and Giovanni Mastrobuoni. Police patrols and crime. *Working Paper*, 2018.
- Guido Imbens and Karthik Kalyanaraman. Optimal Bandwidth Choice for the Regression Discontinuity Estimator. *Review of Economic Studies*, 79(3):933–959, 2012.
- Guido Imbens and Thomas Lemieux. Regression discontinuity designs: A guide to practice. *Journal of Econometrics*, 142(2):615–635, 2008.
- Damon Jones and Ioana Marinescu. The Labor Market Impacts of Universal and Permanent Cash Transfers: Evidence from the Alaska Permanent Fund. NBER Working Papers 24312, National Bureau of Economic Research, Inc, February 2018.

- Jodi Kantor and Megan Twohey. *She said*. New York, Penguin, 2019.
- Jacob Kaplan. Uniform crime reporting program data: Offenses known and clearances by arrest, 1960-2016, 2018.
- Patrick Radden Keefe. The empire of edge. *The New York Times*, October 2014.
- Jonathan Klick and Alexander Tabarrok. Using terror alert levels to estimate the effect of police on crime. *The Journal of Law and Economics*, 48(1):267–279, 2005.
- John Knowles, Nicola Persico, and Petra Todd. Racial bias in motor vehicle searches: Theory and evidence. *Journal of Political Economy*, 109(1):203–229, 2001.
- Frances Xu Lee and Wing Suen. Credibility of crime allegations. *American Economic Journal: Microeconomics*, Forthcoming.
- Steven D Levitt. Using Electoral Cycles in Police Hiring to Estimate the Effect of Police on Crime. *American Economic Review*, 87(3):270–290, June 1997.
- Heather Mac Donald. The war on cops. New York: Encounter Books, 2016.
- John M. MacDonald. De-policing as a consequence of the so-called “ferguson effect”. *Criminology & Public Policy*, 18(1):47–49, 2019.
- Michael D Maltz and Harald E Weiss. Creating a ucr utility: Final report to the national institute of justice. *NIJ Research Report*, 15341:1–21, August 2006.
- Charles F. Manski. Optimal search profiling with linear deterrence. *The American Economic Review*, 95(2):122–126, 2005.
- Charles F. Manski. Search profiling with partial knowledge of deterrence. *The Economic Journal*, 116(515):F385–F401, 2006.
- Charles F Manski and Daniel S Nagin. Assessing benefits, costs, and disparate racial impacts of confrontational proactive policing. *Proceedings of the National Academy of Sciences*, 2017.

- Alexandre Mas. Pay, reference points, and police performance. *The Quarterly Journal of Economics*, 121(3):783–821, 2006.
- Steven Mello. More cops, less crime. *Journal of Public Economics*, 172:174 – 200, 2019.
- I. Durre B. Korzeniewski S. McNeal K. Thomas X. Yin S. Anthony R. Ray R.S. Vose B.E.Gleason Menne, M.J. and T.G. Houston. Global historical climatology network - daily (ghcn-daily). 2012.
- Amalia R Miller and Carmit Segal. Do female officers improve law enforcement quality? effects on crime reporting and domestic violence. *The Review of Economic Studies*, page rdy051, 2018.
- Sara M. Moniuszko. List: All of the hollywood power players accused of sexual assault or harassment. November 2017.
- Jonathan Mummolo. Modern police tactics, police-citizen interactions and the prospects for reform. *The Journal of Politics*, 80(1):1–15, 2018.
- Jerzy S Neyman. On the application of probability theory to agricultural experiments. essay on principles. section 9.(translated and edited by dm dabrowska and tp speed, statistical science (1990), 5, 465-480). *Annals of Agricultural Sciences*, 10:1–51, 1923.
- Emily Owens. Economic approach to "de-policing". *Criminology & Public Policy*, 18(1): 77–80, 2019.
- Emily Owens, David Weisburd, Karen L. Amendola, and Geoffrey P. Alpert. Can you build a better cop? *Criminology & Public Policy*, 17(1):41–87, 2018.
- Harry Pei and Bruno Strulovici. Crime Entanglement, Deterrence, and Witness Credibility. *Working Paper*, 2019.
- Nicola Persico. Racial profiling, fairness, and effectiveness of policing. *The American Economic Review*, 92(5):1472–1497, 2002.
- Andrea Prat. The wrong kind of transparency. *The American Economic Review*, 95(3):862–877, 2005.

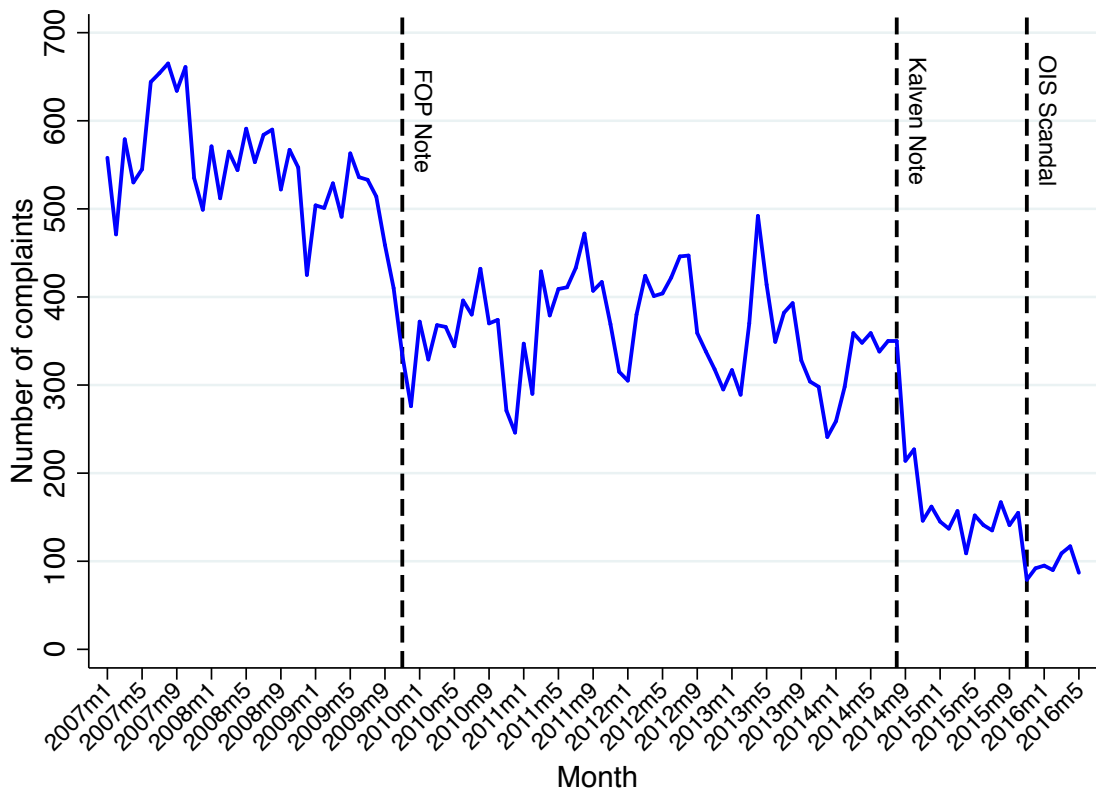
- Canice Prendergast. Selection and oversight in the public sector, with the los angeles police department as an example. Technical report, National Bureau of Economic Research, 2001.
- Roni Caryn Rabin. Doctors behaving badly. *New York Times*, August 2015.
- Kyle Rozema and Max Schanzenbach. Good cop, bad cop: Using civilian allegations to predict police misconduct. *American Economic Journal: Economic Policy*, forthcoming, 2018.
- Donald B Rubin. Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of educational Psychology*, 66(5):688, 1974.
- Joanna C Schwartz. Police indemnification. 2014.
- Joanna C Schwartz. How governments pay: Lawsuits, budgets, and police reform. *UCLA L. Rev.*, 63:1144, 2016.
- Lan Shi. The limit of oversight in policing: Evidence from the 2001 cincinnati riot. *Journal of Public Economics*, 93(1):99–113, 2009.
- Wesley Skogan and Kathleen Frydl. *Fairness and Effectiveness in Policing: The Evidence*. The National Academies Press, Washington, DC, 2004.
- Alexandra Stevenson and Matthew Goldstein. At world’s largest hedge fund, sex, fear and video surveillance. *DealBook, The New York Times*, July 2016.
- Christopher Stone, Todd S Foglesong, and Christine M Cole. *Policing Los Angeles under a consent degree: The dynamics of change at the LAPD*. Program in Criminal Justice Policy and Management, Harvard Kennedy School, 2009.
- David R. Thompson and Alexander M. Clark. The ego has landed! what can be done about research misconduct, scandals, and spins? *The American Journal of Medicine*, 132(5):552–553, 2019.

Tom R. Tyler, Phillip Atiba Goff, and Robert J. MacCoun. The impact of psychological science on policing in the united states: Procedural justice, legitimacy, and effective law enforcement. *Psychological Science in the Public Interest*, 16(3):75–109, 2015.

Sarit Weisburd. Police presence, rapid response rates, and crime prevention. *Working Paper*, 2018.

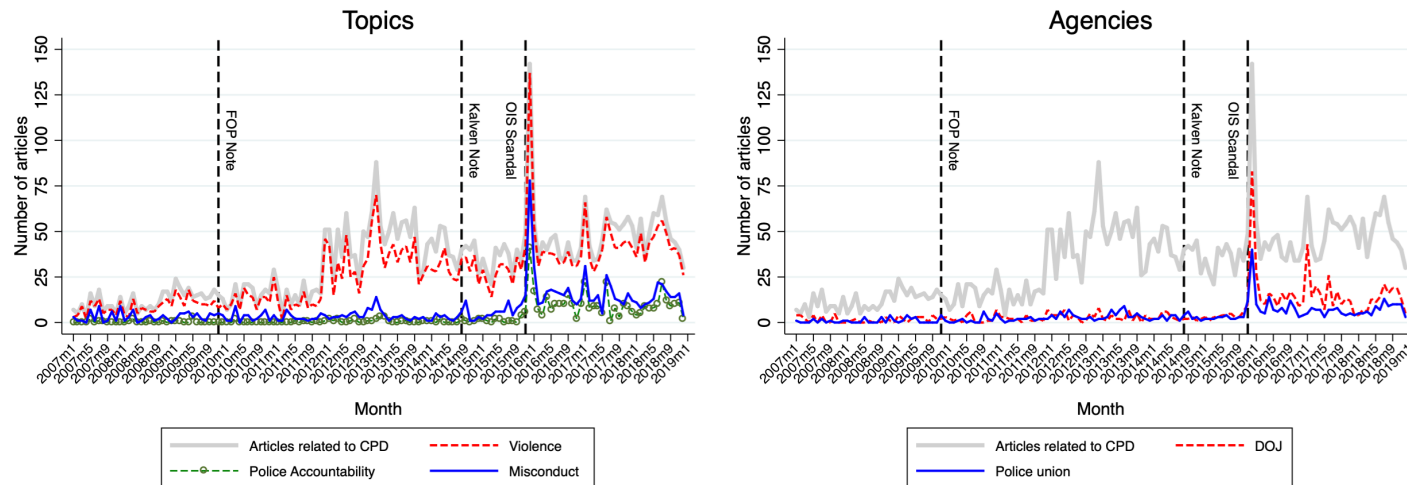
Yiqing Xu. Generalized synthetic control method: Causal inference with interactive fixed effects models. *Political Analysis*, 25(1):57–76, 2017.

Figure 1: Complaints from 2007-2016



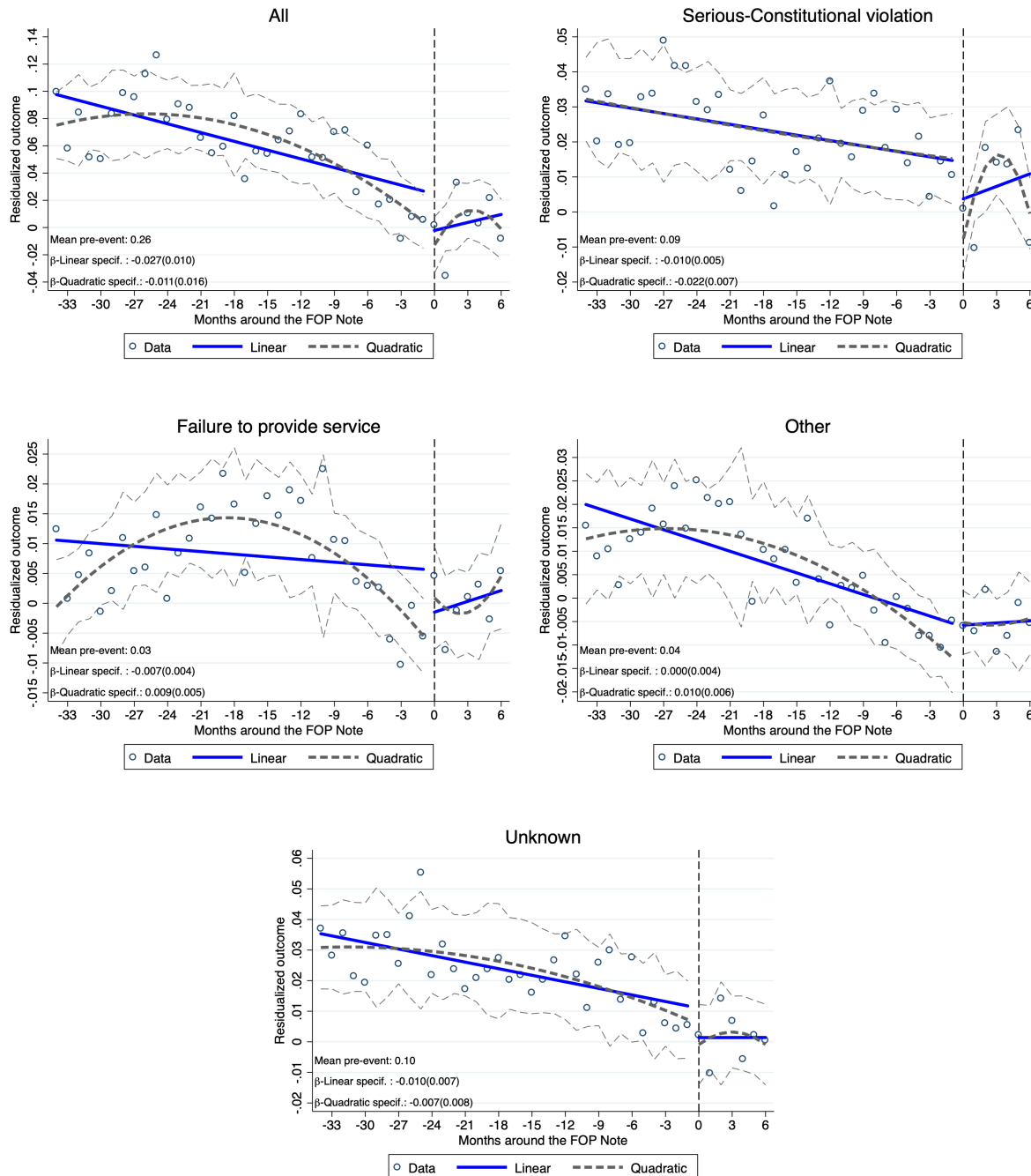
Notes: This figure presents the total number of allegation of misconduct against a Chicago police officer from January 2007 to May 2016. The dashed lines represent three events that have varied the level of monitoring against Chicago Police Department between 2007 and 2016: a note from the FOP, a notification from the police union stating that complaints against Chicago police officers are available to the public (Kalven Note), and a scandal related to an officer shooting and killing an adolescent (OIS Scandal). Section 2, in the main text, describes each event in detail.

Figure 2: Articles related to the Chicago Police Department in the Chicago Tribune



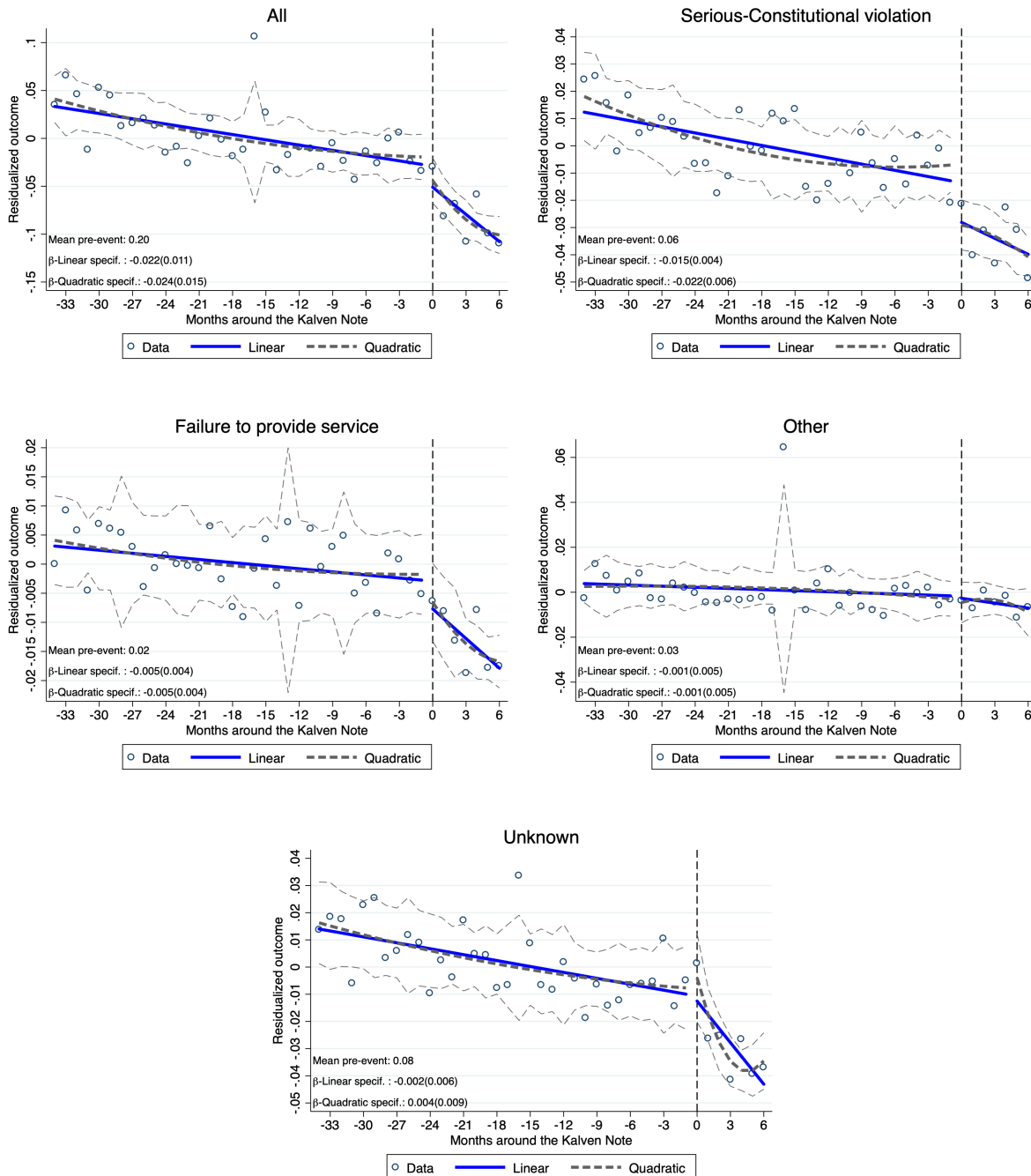
Notes: These figures present the number of articles mentioning the Chicago Police Department in the Chicago Tribune from January 2007 to December 2018. The left-figure presents the frequency distribution of articles related to gun-violence, police accountability, and police misconduct. The left-figure presents the frequency distribution of articles mentioning the DOJ and the Chicago police union.

Figure 3: Effect of the FOP Note on the number of complaints per 10,000 residents



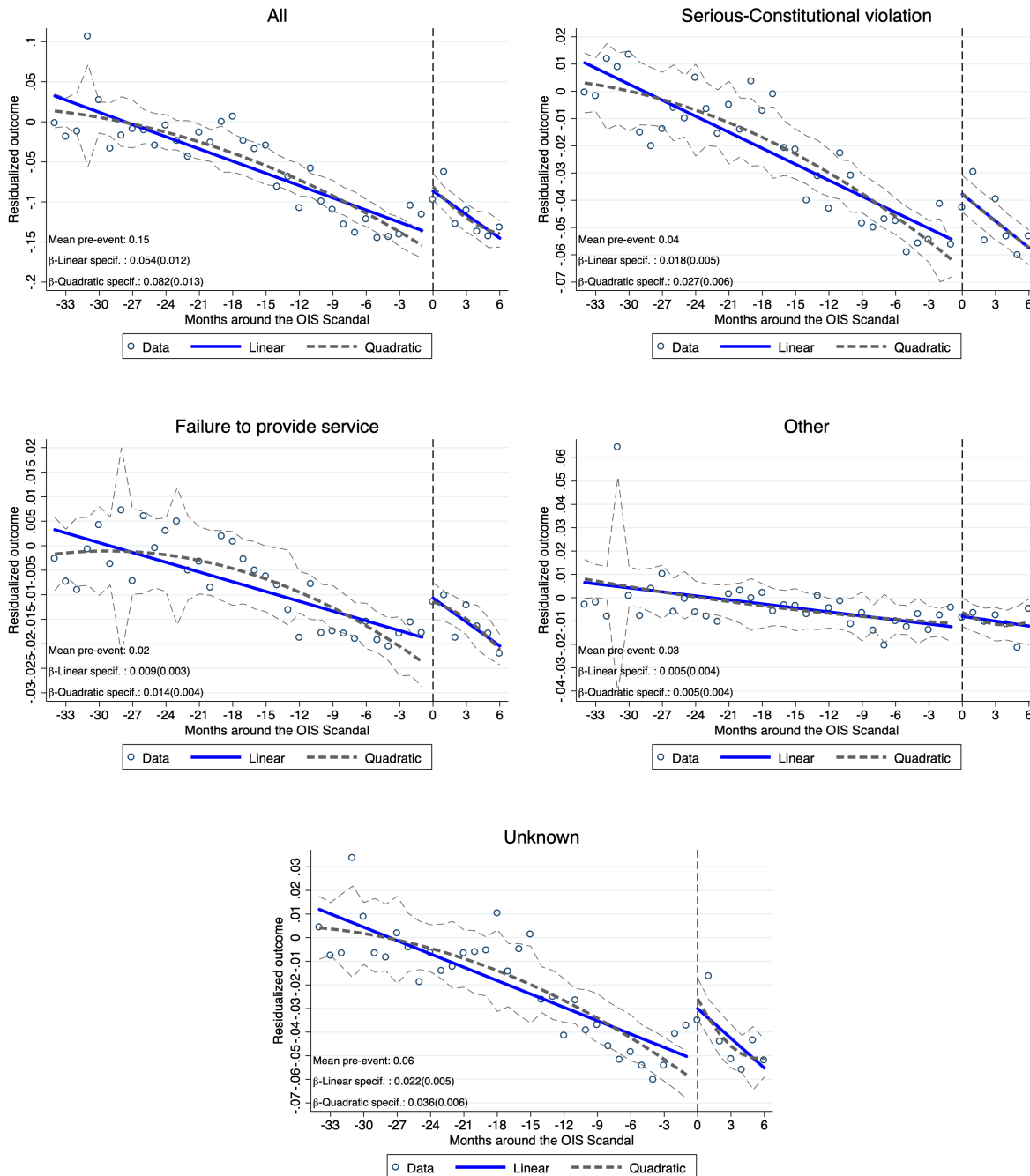
Notes: These figures present the impact of the FOP Notes on complaints per 10,000 residents. We report the lines from local linear and quadratic regressions, estimated separately on each side of the cutoff. We account for seasonal patterns in outcomes by removing the month effect and report the residuals of this regression. We report the mean of the dependent variable and RD estimates. Standard errors are clustered at the police beat level and are reported in parentheses. Dashed lines correspond to the 90% confidence interval for the quadratic specification. In each graph, the scale of the y-axis is set equal to ± 5 standard deviation of the respective variable.

Figure 4: Effect of the Kalven Note on the number of complaints per 10,000 residents



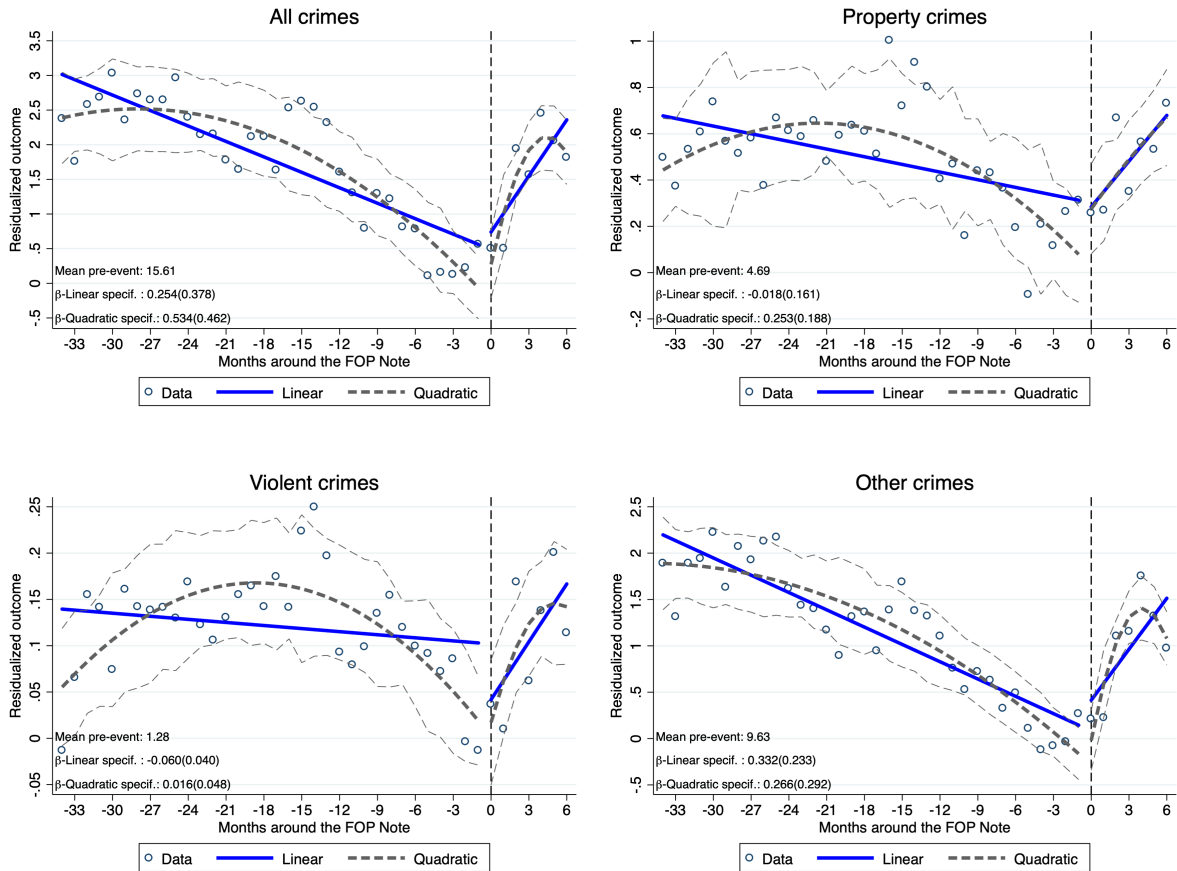
Notes: These figures present the impact of the Kalven Note on complaints per 10,000 residents. We report the lines from local linear and quadratic regressions, estimated separately on each side of the cutoff. We account for seasonal patterns in outcomes by removing the month effect and report the residuals of this regression. We report the mean of the dependent variable and RD estimates. Standard errors are clustered at the police beat level and are reported in parentheses. Dashed lines correspond to the 90% confidence interval for the quadratic specification. In each graph, the scale of the y-axis is set equal to ± 5 standard deviation of the respective variable.

Figure 5: Effect of the OIS Scandal on the number of complaints per 10,000 residents



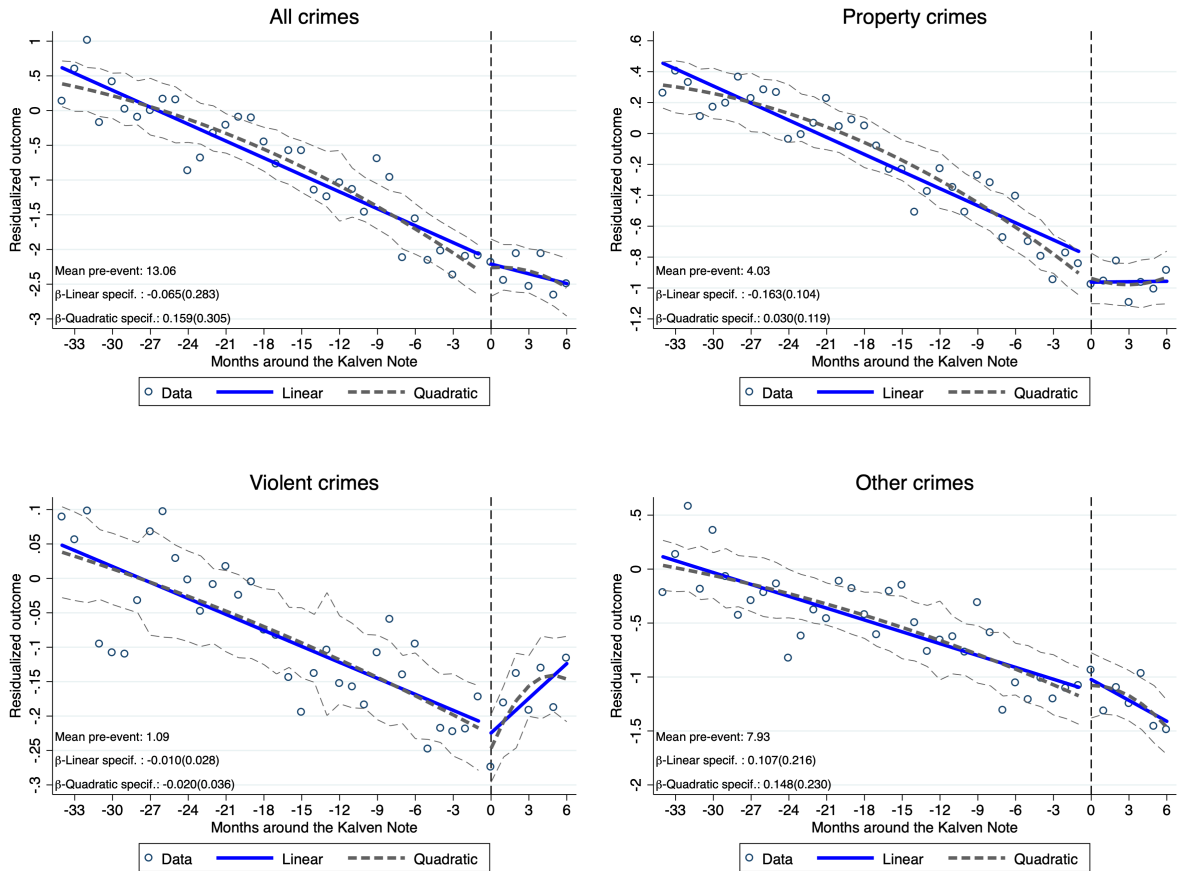
Notes: These figures present the impact of the OIS Scandal on complaints per 10,000 residents. We report the lines from local linear and quadratic regressions, estimated separately on each side of the cutoff. We account for seasonal patterns in outcomes by removing the month effect and report the residuals of this regression. We report the mean of the dependent variable and RD estimates. Standard errors are clustered at the police beat level and are reported in parentheses. Dashed lines correspond to the 90% confidence interval for the quadratic specification. In each graph, the scale of the y-axis is set equal to ± 0.5 standard deviation of the respective variable.

Figure 6: Effect of the FOP Note on the number of crimes per 10,000 residents



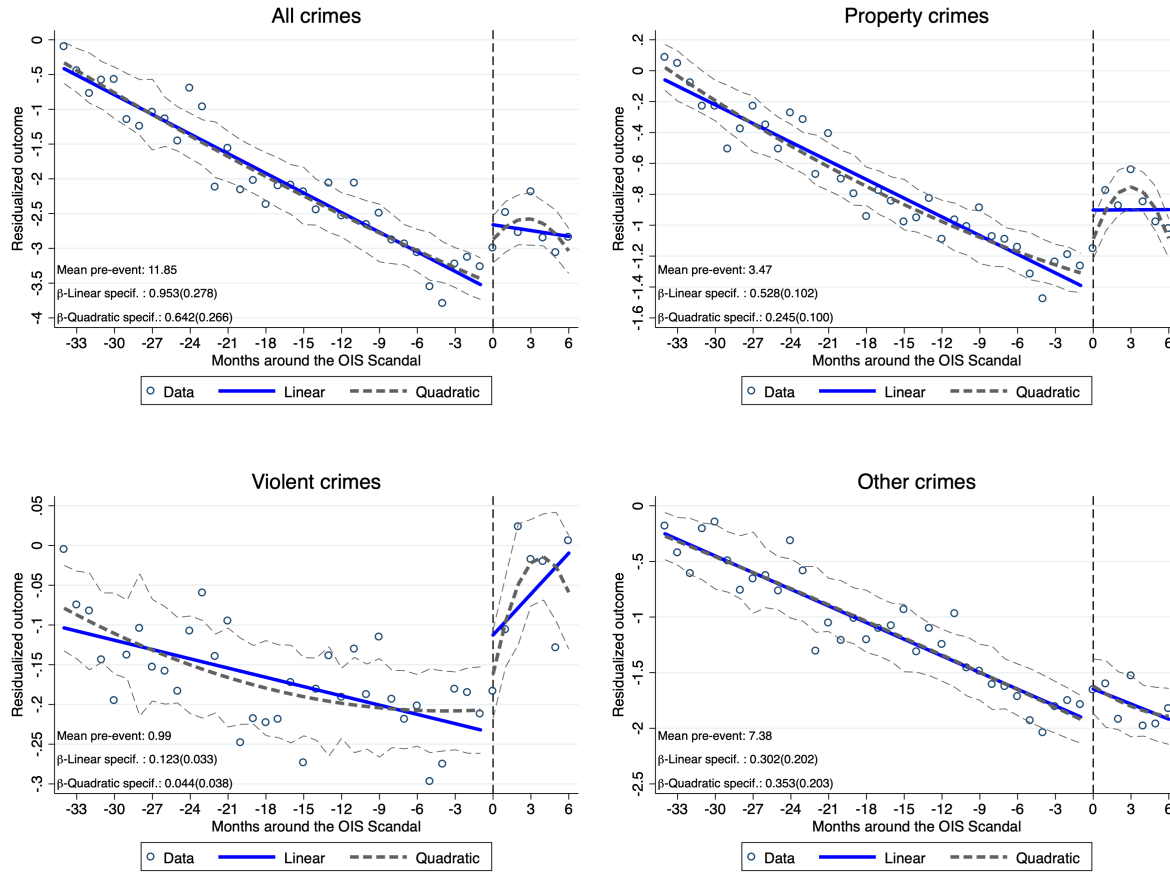
Notes: These figures present the impact of the FOP Note on crimes per 10,000 residents. We report the lines from local linear and quadratic regressions, estimated separately on each side of the cutoff. We account for seasonal patterns in outcomes by removing the month effect and report the residuals of this regression. We report the mean of the dependent variable and RD estimates. Standard errors are clustered at the police beat level and are reported in parentheses. Dashed lines correspond to the 90% confidence interval for the quadratic specification. In each graph, the scale of the y-axis is set equal to ± 0.5 standard deviation of the respective variable.

Figure 7: Effect of the Kalven Note on the number of crimes per 10,000 residents



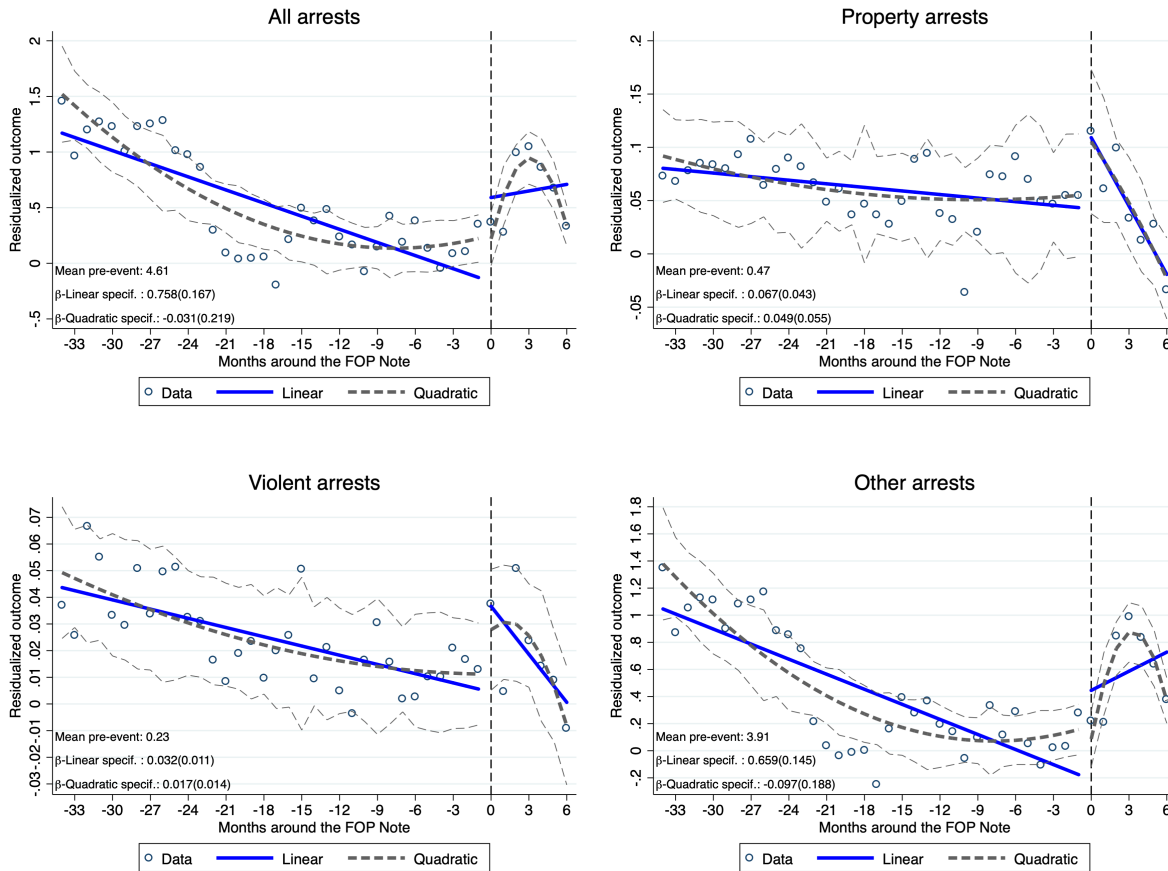
Notes: These figures present the impact of the Kalven Note on crimes per 10,000 residents. We report the lines from local linear and quadratic regressions, estimated separately on each side of the cutoff. We account for seasonal patterns in outcomes by removing the month effect and report the residuals of this regression. We report the mean of the dependent variable and RD estimates. Standard errors are clustered at the police beat level and are reported in parentheses. Dashed lines correspond to the 90% confidence interval for the quadratic specification. In each graph, the scale of the y-axis is set equal to ± 0.5 standard deviation of the respective variable.

Figure 8: Effect of the OIS Scandal on the number of crimes per 10,000 residents



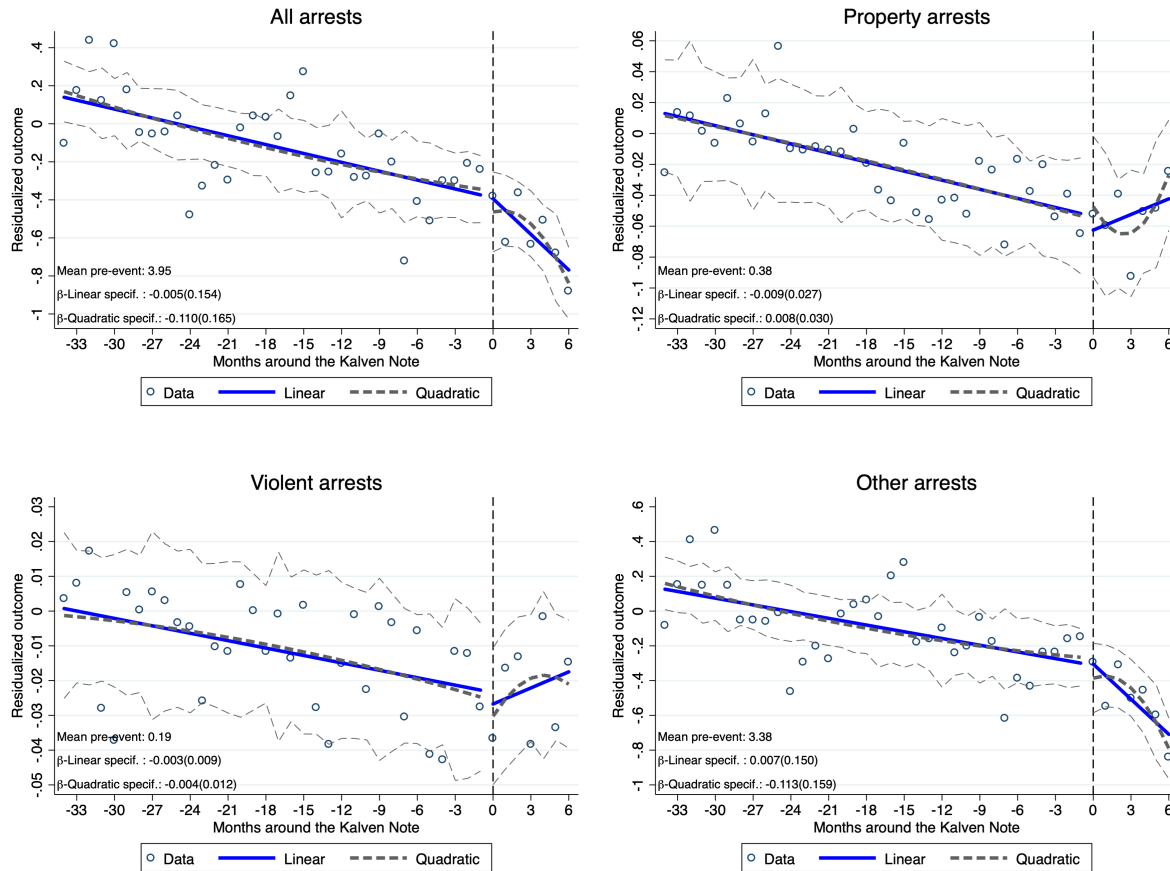
Notes: These figures present the impact of the OIS Scandal on crimes per 10,000 residents. We report the lines from local linear and quadratic regressions, estimated separately on each side of the cutoff. We account for seasonal patterns in outcomes by removing the month effect and report the residuals of this regression. We report the mean of the dependent variable and RD estimates. Standard errors are clustered at the police beat level and are reported in parentheses. Dashed lines correspond to the 90% confidence interval for the quadratic specification. In each graph, the scale of the y-axis is set equal to ± 5 standard deviation of the respective variable.

Figure 9: Effect of the FOP Note on the number of arrests per 10,000 residents



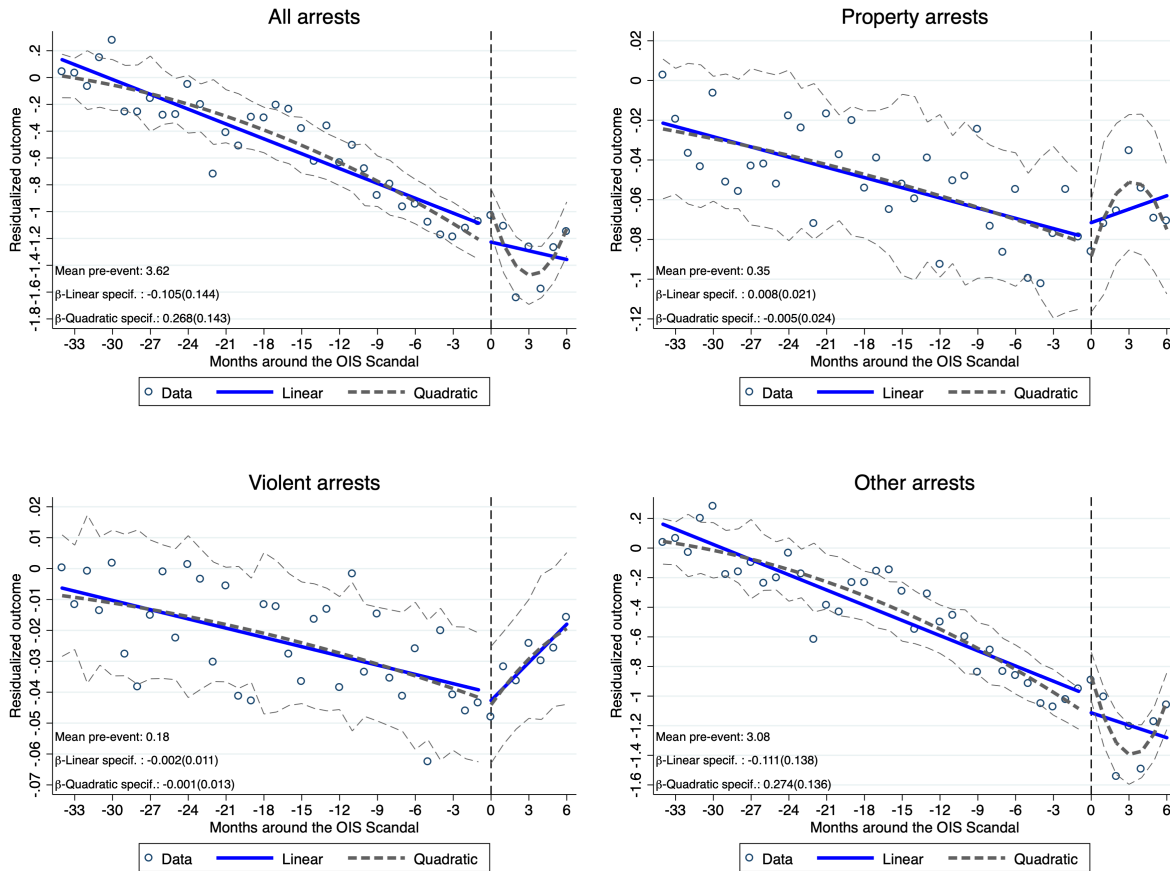
Notes: These figures present the impact of the FOP Note on arrests per 10,000 residents. We report the lines from local linear and quadratic regressions, estimated separately on each side of the cutoff. We account for seasonal patterns in outcomes by removing the month effect and report the residuals of this regression. We report the mean of the dependent variable and RD estimates. Standard errors are clustered at the police beat level and are reported in parentheses. Dashed lines correspond to the 90% confidence interval for the quadratic specification. In each graph, the scale of the y-axis is set equal to ± 5 standard deviation of the respective variable.

Figure 10: Effect of the Kalven Note on the number of arrests per 10,000 residents



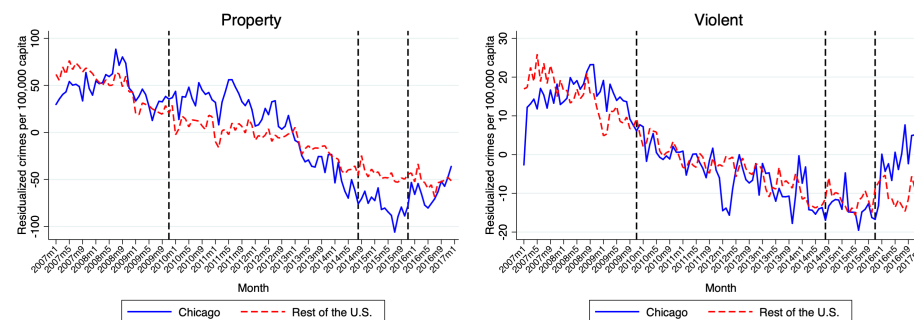
Notes: These figures present the impact of the Kalven Note on arrests per 10,000 residents. We report the lines from local linear and quadratic regressions, estimated separately on each side of the cutoff. We account for seasonal patterns in outcomes by removing the month effect and report the residuals of this regression. We report the mean of the dependent variable and RD estimates. Standard errors are clustered at the police beat level and are reported in parentheses. Dashed lines correspond to the 90% confidence interval for the quadratic specification. In each graph, the scale of the y-axis is set equal to ± 5 standard deviation of the respective variable.

Figure 11: Effect of the OIS Scandal on the number of arrests per 10,000 residents



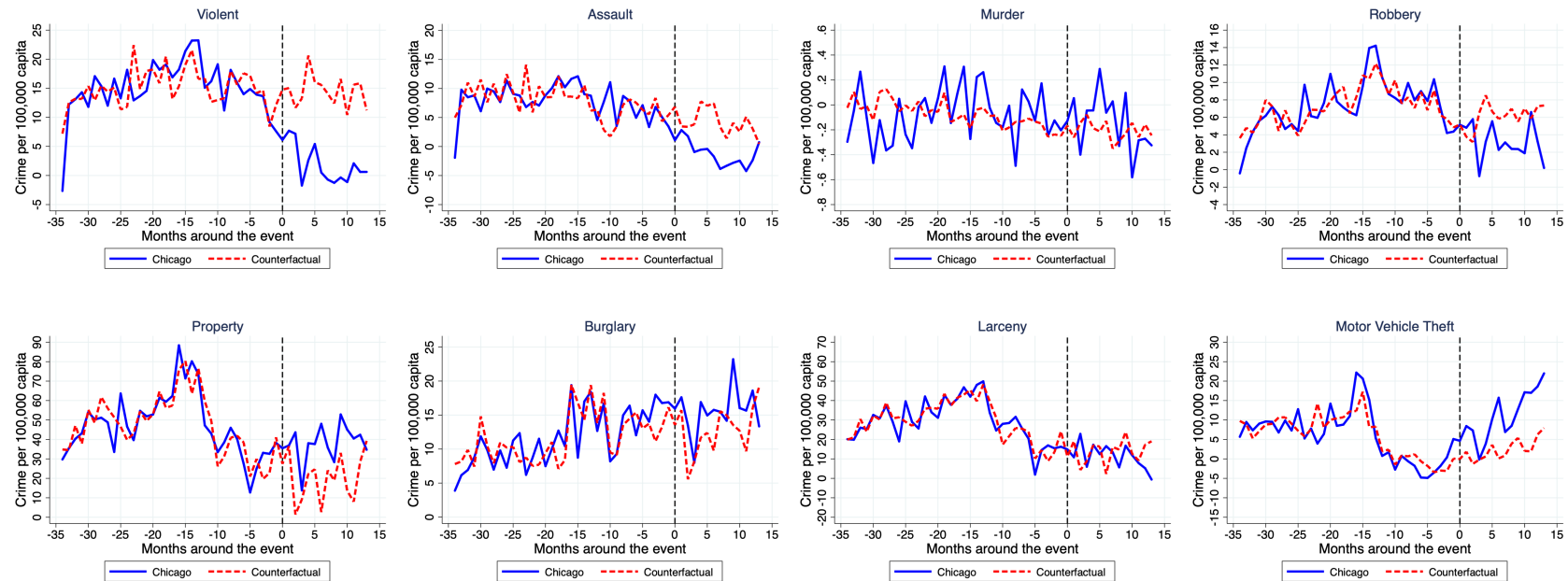
Notes: These figures present the impact of the OIS Scandal on arrests per 10,000 residents. We report the lines from local linear and quadratic regressions, estimated separately on each side of the cutoff. We account for seasonal patterns in outcomes by removing the month effect and report the residuals of this regression. We report the mean of the dependent variable and RD estimates. Standard errors are clustered at the police beat level and are reported in parentheses. Dashed lines correspond to the 90% confidence interval for the quadratic specification. In each graph, the scale of the y-axis is set equal to ± 0.5 standard deviation of the respective variable.

Figure 12: Seasonality adjusted crimes rates: Chicago vs. the rest of the U.S.



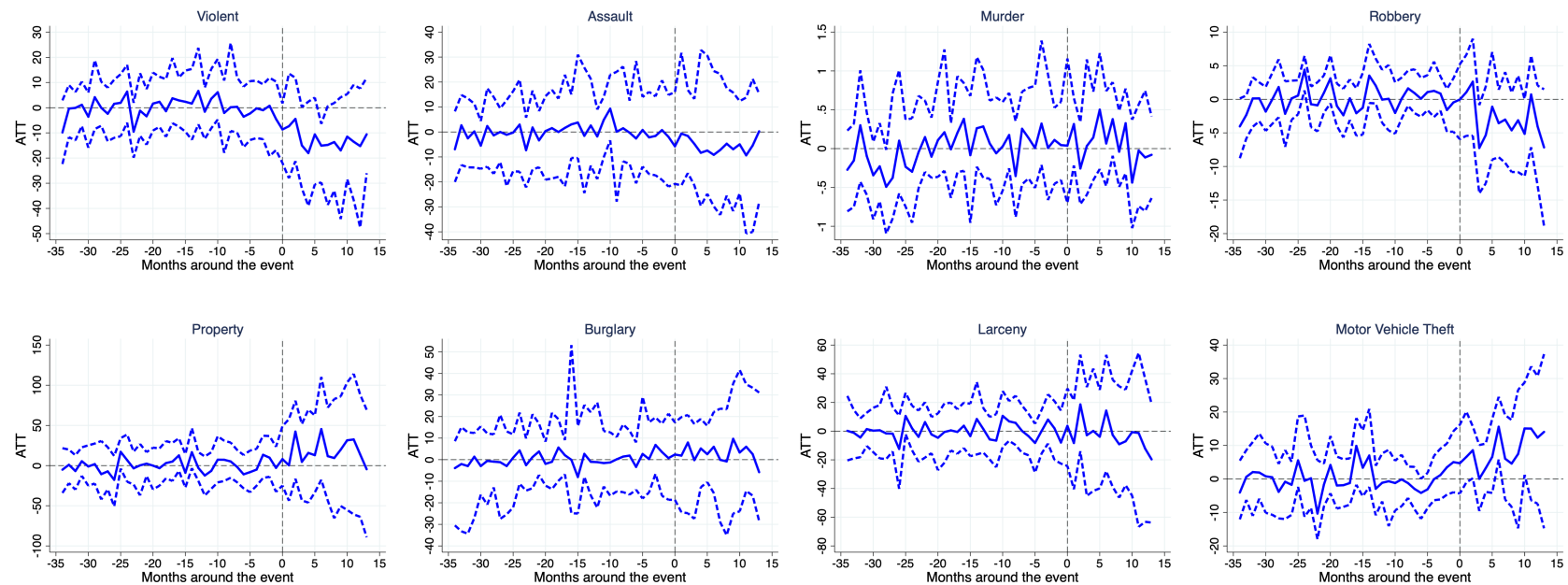
Notes: These figures present the number of reported property and violent crimes per 100,000 capita in Chicago and the rest of U.S. from January 2007 to December 2016. We account for seasonal patterns in crime outcomes, i.e., we remove the month effect and report the residuals of this regression (some outcomes can have negative values). The vertical dashed lines represent three events that have varied the level of monitoring against Chicago Police Department between 2007 and 2016 (see Section 2).

Figure 13: Trends in crimes before and after the FOP Note using generalized synthetic control methods: treated vs. counterfactual



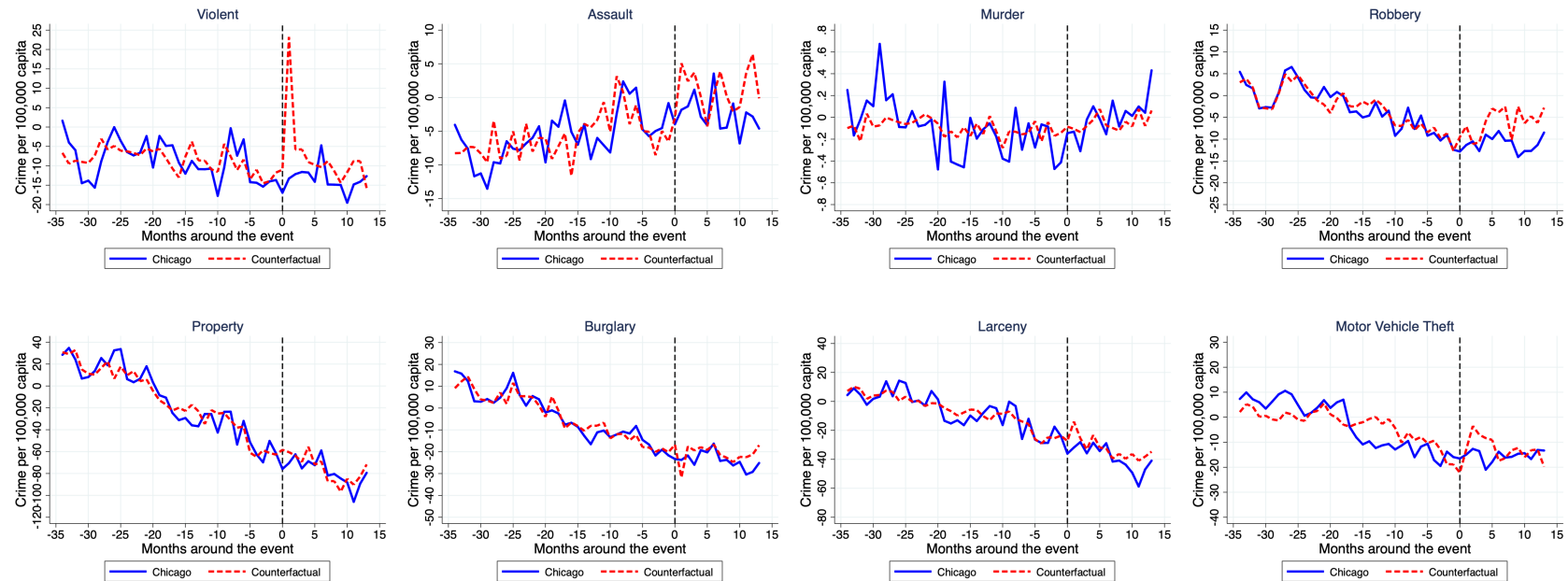
Notes: These figures present the trends in crimes per 100,000 capita in Chicago and its counterfactual from January 2007 to December 2016. We account for seasonal patterns in crime outcomes, i.e., we remove the month effect and report the residuals of this regression (some outcomes can have negative values). The vertical dashed line represent the time of the event that varied the level of monitoring against Chicago Police Department between 2007 and 2016 (see Section 2). The event of interest corresponds to the FOP Note. In each graph, the scale of the y-axis is set equal to ± 0.5 standard deviation of the respective variable.

Figure 14: Estimated Average Treatment Effect before and after the FOP Note using generalized synthetic control methods



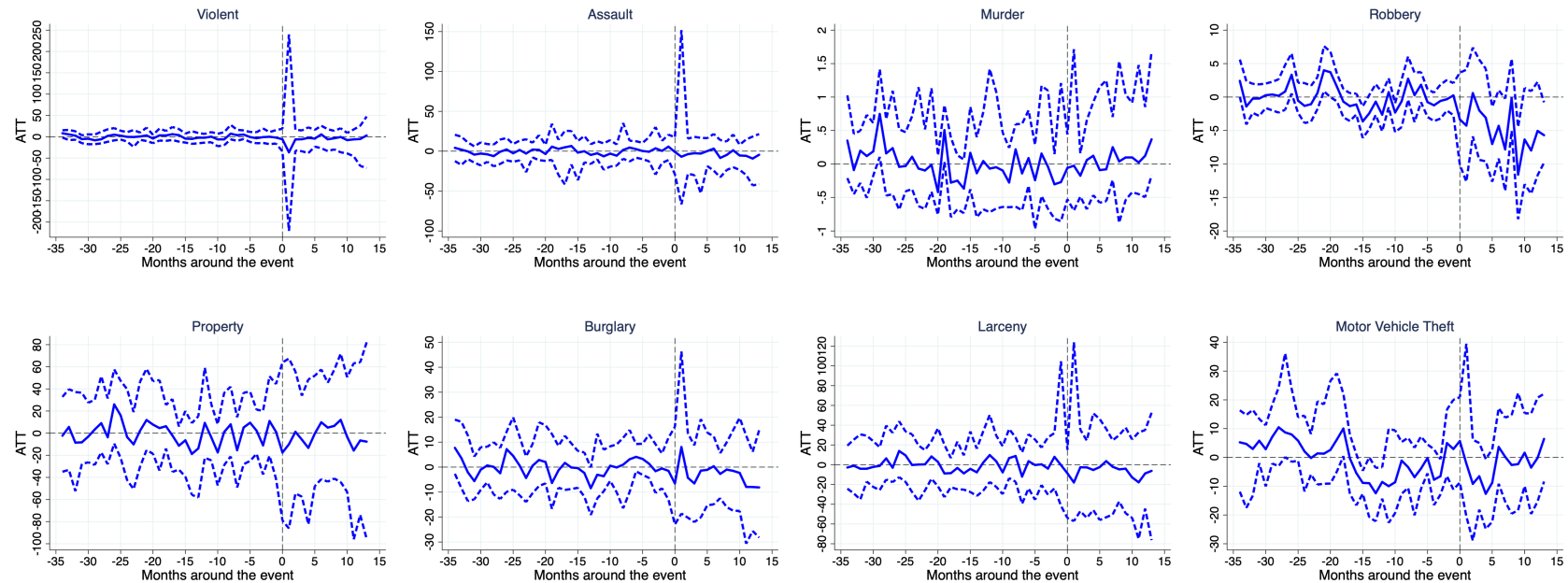
Notes: These figures present the estimated average treatment effect of the event on the number of crimes per 100,000 capita in Chicago and its counterfactual from January 2007 to December 2016. Dashed lines correspond to the 95% confidence interval. We account for seasonal patterns in crime outcomes, i.e., we remove the month effect and report the residuals of this regression (some outcomes can have negative values). The vertical dashed line represent the time of the event that varied the level of monitoring against Chicago Police Department between 2007 and 2016 (see Section 2). The event of interest corresponds to the FOP Note. In each graph, the scale of the y-axis is set equal to ± 0.5 standard deviation of the respective variable.

Figure 15: Trends in crimes before and after the Kalven Note using generalized synthetic control methods: treated vs. counterfactual



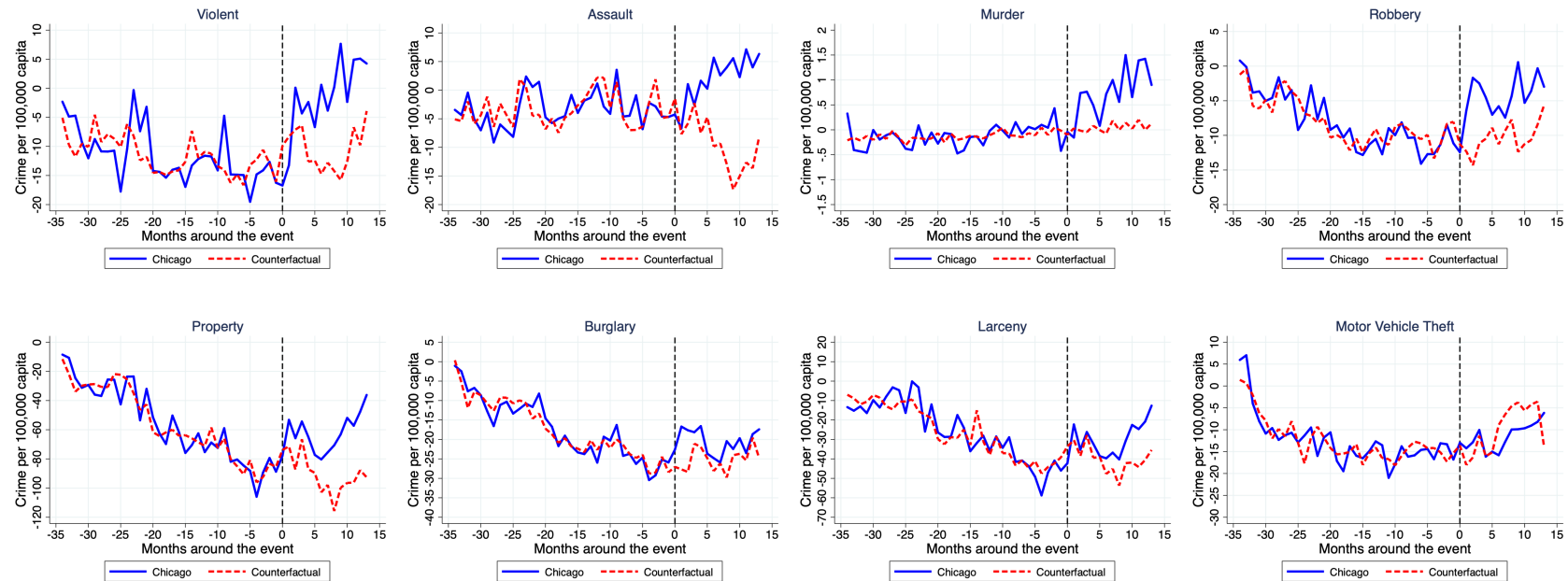
Notes: These figures present the trends in crimes per 100,000 capita in Chicago and its counterfactual from January 2007 to December 2016. We account for seasonal patterns in crime outcomes, i.e., we remove the month effect and report the residuals of this regression (some outcomes can have negative values). The vertical dashed line represent the time of the event that varied the level of monitoring against Chicago Police Department between 2007 and 2016 (see Section 2). The event of interest corresponds to the Kalven Note. In each graph, the scale of the y-axis is set equal to ± 0.5 standard deviation of the respective variable.

Figure 16: Estimated Average Treatment Effect in crimes before and after the Kalven Note using generalized synthetic control methods



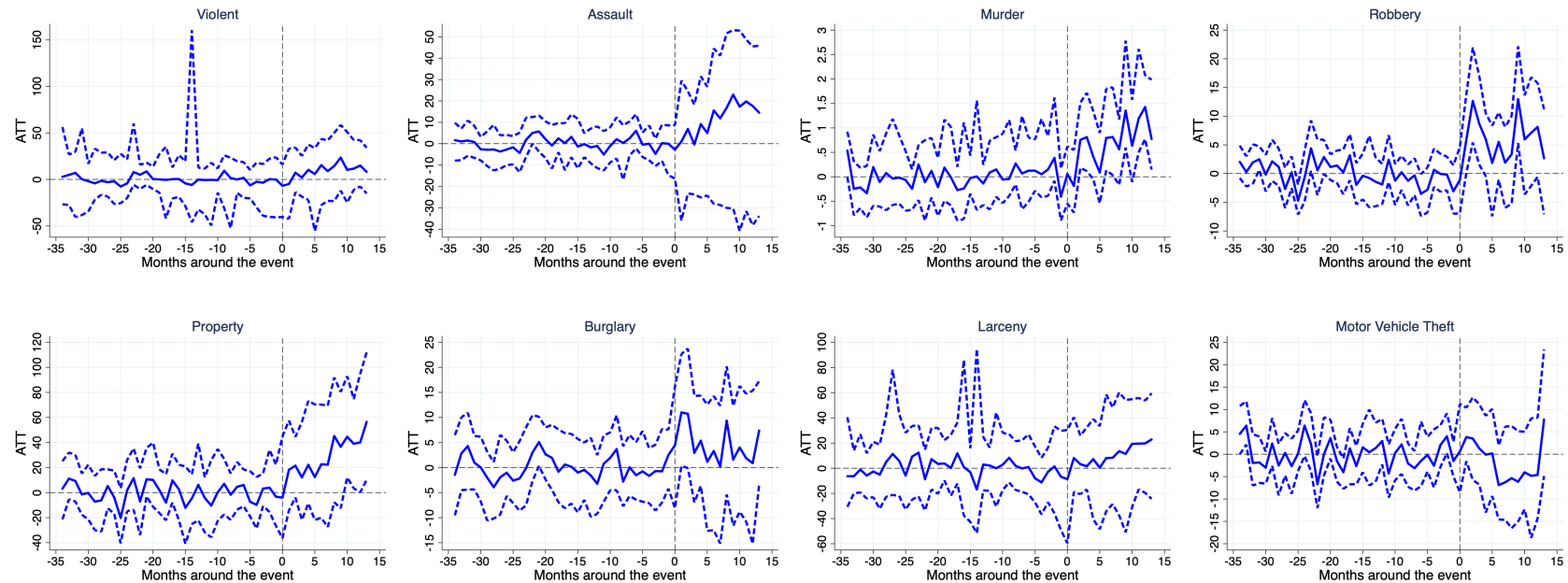
Notes: These figures present the estimated average treatment effect of the event on the number of crimes per 100,000 capita in Chicago and its counterfactual from January 2007 to December 2016. Dashed lines correspond to the 95% confidence interval. We account for seasonal patterns in crime outcomes, i.e., we remove the month effect and report the residuals of this regression (some outcomes can have negative values). The vertical dashed line represent the time of the event that varied the level of monitoring against Chicago Police Department between 2007 and 2016 (see Section 2). The event of interest corresponds to the Kalven Note. In each graph, the scale of the y-axis is set equal to ± 0.5 standard deviation of the respective variable.

Figure 17: Trends in crimes before and after the OIS Scandal using generalized synthetic control methods: treated vs. counterfactual



Notes: These figures present the trends in crimes per 100,000 capita in Chicago and its counterfactual from January 2007 to December 2016. We account for seasonal patterns in crime outcomes, i.e., we remove the month effect and report the residuals of this regression (some outcomes can have negative values). The vertical dashed line represent the time of the event that varied the level of monitoring against Chicago Police Department between 2007 and 2016 (see Section 2). The event of interest corresponds to the OIS Scandal. In each graph, the scale of the y-axis is set equal to ± 5 standard deviation of the respective variable.

Figure 18: Estimated Average Treatment Effect in crimes before and after the OIS Scandal using generalized synthetic control methods



Notes: These figures present the estimated average treatment effect of the event on the number of crimes per 100,000 capita in Chicago and its counterfactual from January 2007 to December 2016. Dashed lines correspond to the 95% confidence interval. We account for seasonal patterns in crime outcomes, i.e., we remove the month effect and report the residuals of this regression (some outcomes can have negative values). The vertical dashed line represent the time of the event that varied the level of monitoring against Chicago Police Department between 2007 and 2016 (see Section 2). The event of interest corresponds to the OIS Scandal. In each graph, the scale of the y-axis is set equal to ± 0.5 standard deviation of the respective variable.

Table 1: Complaints response to monitoring

	All (1)	All (2)	Serious (3)	Serious (4)	FPS (5)	FPS (6)	Other (7)	Other (8)	Unkn. (9)	Unkn. (10)
Panel A, FOP Note										
Event	-0.027** (0.010)	-0.011 (0.016)	-0.010** (0.005)	-0.022*** (0.007)	-0.007** (0.004)	0.009* (0.005)	0.000 (0.004)	0.010 (0.006)	-0.010 (0.007)	-0.007 (0.008)
Mean pre-event	0.259	0.259	0.088	0.088	0.034	0.034	0.040	0.040	0.098	0.098
Panel B, Kalven Note										
Event	-0.022* (0.011)	-0.024* (0.015)	-0.015*** (0.004)	-0.022*** (0.006)	-0.005 (0.004)	-0.005 (0.004)	-0.001 (0.005)	-0.001 (0.005)	-0.002 (0.006)	0.004 (0.009)
Mean pre-event	0.197	0.197	0.063	0.063	0.025	0.025	0.033	0.033	0.076	0.076
Panel C, OIS Scandal										
Event	0.054*** (0.012)	0.082*** (0.013)	0.018*** (0.005)	0.027*** (0.006)	0.009*** (0.003)	0.014*** (0.004)	0.005 (0.004)	0.005 (0.004)	0.022*** (0.005)	0.036*** (0.006)
Mean pre-event	0.147	0.147	0.044	0.044	0.018	0.018	0.029	0.029	0.056	0.056
No of obs.	9384	9384	9384	9384	9384	9384	9384	9384	9384	9384
No of beats	276	276	276	276	276	276	276	276	276	276
Specification	Linear	Quadr.	Linear	Quadr.	Linear	Quadr.	Linear	Quadr.	Linear	Quadr.

Notes: This Table presents the effect of oversights on the the number of complaints per 10,000 residents in Chicago from January 2007 to May 2016 (see Section 2). We account for seasonal patterns in outcomes by removing the month effect and report the residuals of this regression. We report the mean of the dependent variable at pre-event and RD estimates. Standard errors are clustered at the police beat level and are reported in parentheses. *p-value < 0.10, **p-value < 0.05, *** p-value < 0.01.

Table 2: Crimes response to monitoring

	All Crimes (1)	All Crimes (2)	Violent Crimes (3)	Violent Crimes (4)	Property Crimes (5)	Property Crimes (6)	Other Crimes (7)	Other Crimes (8)
Panel A, FOP Note								
Event	0.254 (0.378)	0.534 (0.462)	-0.060 (0.040)	0.016 (0.048)	-0.018 (0.161)	0.253 (0.188)	0.332 (0.233)	0.266 (0.292)
Mean pre-event	15.608	15.608	1.285	1.285	4.693	4.693	9.630	9.630
Panel B, Kalven Note								
Event	-0.065 (0.283)	0.159 (0.305)	-0.010 (0.028)	-0.020 (0.036)	-0.163 (0.104)	0.030 (0.119)	0.107 (0.216)	0.148 (0.230)
Mean pre-event	13.056	13.056	1.088	1.088	4.034	4.034	7.933	7.933
Panel C, OIS Scandal								
Event	0.953*** (0.278)	0.642** (0.266)	0.123*** (0.033)	0.044 (0.038)	0.528*** (0.102)	0.245** (0.100)	0.302 (0.202)	0.353* (0.203)
Mean pre-event	11.845	11.845	0.994	0.994	3.468	3.468	7.383	7.383
No of obs.	9384	9384	9384	9384	9384	9384	9384	9384
No of beats	276	276	276	276	276	276	276	276
Specification	Linear	Quadr.	Linear	Quadr.	Linear	Quadr.	Linear	Quadr.

Notes: This Table presents the effect of oversights on the the number of crimes per 10,000 residents in Chicago from January 2007 to May 2016 (see Section 2). We account for seasonal patterns in outcomes by removing the month effect and report the residuals of this regression. We report the mean of the dependent variable at pre-event and RD estimates. Standard errors are clustered at the police beat level and are reported in parentheses. *p-value < 0.10, **p-value < 0.05, *** p-value < 0.01.

Table 3: Arrests response to monitoring

	All Arrests (1)	All Arrests (2)	Violent Arrests (3)	Violent Arrests (4)	Property Arrests (5)	Property Arrests (6)	Other Arrests (7)	Other Arrests (8)
Panel A, FOP Note								
Event	0.758*** (0.167)	-0.031 (0.219)	0.032*** (0.011)	0.017 (0.014)	0.067 (0.043)	0.049 (0.055)	0.659*** (0.145)	-0.097 (0.188)
Mean pre-event	4.607	4.607	0.229	0.229	0.467	0.467	3.911	3.911
Panel B, Kalven Note								
Event	-0.005 (0.154)	-0.110 (0.165)	-0.003 (0.009)	-0.004 (0.012)	-0.009 (0.027)	0.008 (0.030)	0.007 (0.150)	-0.113 (0.159)
Mean pre-event	3.952	3.952	0.194	0.194	0.381	0.381	3.377	3.377
Panel C, OIS Scandal								
Event	-0.105 (0.144)	0.268* (0.143)	-0.002 (0.011)	-0.001 (0.013)	0.008 (0.021)	-0.005 (0.024)	-0.111 (0.138)	0.274** (0.136)
Mean pre-event	3.619	3.619	0.183	0.183	0.352	0.352	3.084	3.084
No of obs.	9384	9384	9384	9384	9384	9384	9384	9384
No of beats	276	276	276	276	276	276	276	276
Specification	Linear	Quadr.	Linear	Quadr.	Linear	Quadr.	Linear	Quadr.

Notes: This Table presents the effect of oversights on the the number of arrests per 10,000 residents in Chicago from January 2007 to May 2016 (see Section 2). We account for seasonal patterns in outcomes by removing the month effect and report the residuals of this regression. We report the mean of the dependent variable at pre-event and RD estimates. Standard errors are clustered at the police beat level and are reported in parentheses. *p-value < 0.10, **p-value < 0.05, *** p-value < 0.01.

Table 4: Summary statistics: Chicago vs. the rest of the U.S.

	Rest of U.S.	Chicago
Violent crime	172.29	84.60
Assault	148.94	42.98
Murder	0.75	1.47
Robbery	22.60	40.16
Property crime	351.17	326.20
Burglary	77.44	64.23
Larceny	230.46	215.73
Motor vehicle theft	43.27	46.24
Temperature	61.20	51.94
Precipitation	22.21	29.47
Population	949,658.16	2,750,640
Observations	2880	120

Notes: The sample used is Chicago and a set of 24 large U.S. cities (population > 499,999) from January 2007 to December 2016. The reported crimes are per 100,000 capita and this does not adjust for the seasonality patterns in crime.

Table 5: Effect of monitoring on crime

	<u>FOP Note</u>	<u>Kalven Note</u>	<u>OIS Scandal</u>
	(1)	(2)	(3)
Violent	-12.532** (6.058)	-5.885 (9.59)	8.61 (12.386)
Mean Chicago _{t< t₀}	98.854	75.806	71.793
Assault	-5.428 (8.046)	-3.618 (8.527)	11.099 (19.377)
Mean Chicago _{t< t₀}	50.206	37.096	38.673
Murder	0.061 (0.156)	0.053 (0.156)	0.678*** (0.19)
Mean Chicago _{t< t₀}	1.397	1.353	1.341
Robbery	-2.858 (2.688)	-4.861** (2.012)	5.825* (3.2)
Mean Chicago _{t< t₀}	47.251	37.358	31.779
Property	17.387 (30.862)	-3.272 (22.147)	27.773** (15.095)
Mean Chicago _{t< t₀}	373.701	311.613	269.479
Burglary	2.9 (8.601)	-3.075 (5.043)	4.586 (3.788)
Mean Chicago _{t< t₀}	75.79	60.172	46.479
Larceny	-2.204 (19.347)	-6.617 (15.869)	9.594 (13.925)
Mean Chicago _{t< t₀}	245.693	208.228	189.426
Motor Vehicle Theft	8.57* (4.568)	-2.173 (5.081)	-1.51 (3.67)
Mean Chicago _{t< t₀}	52.219	43.213	33.574
No. of Cities	25	25	25
No. of Observation	1200	1200	1200

Notes: This Table presents the effect of oversight on crimes per 100,000 capita (see Section 2). We account for seasonal patterns in crime outcomes by removing the city-month effect and report the residuals of this regression. Standard errors for the generalized synthetic controls methods (GSC) are based on parametric bootstraps (blocked at the agency level) of 10,000 times. *p-value < 0.10, **p-value < 0.05, *** p-value < 0.01.

Appendix

A Data Appendix

A.1 Cleaning UCR data

The annual city population reported in the FBI files tends to jump discretely around census years (Chalfin and McCrary [2018], Mello [2019]). We follow Mello [2019] and replace the reported population with a smoothed version. For each city, we fit the population time series using local linear regression and replace the reported population with the fitted values.

Similar to Mello [2019], we developed a regression-based procedure to detect cities with extreme outliers and record errors. For each agency, we fit the time series of police reported crimes for assault, burglary, motor vehicle theft, larceny, and robbery using a local linear regression with bandwidth two. We do not use this procedure for murder because it is a relatively rare event. For each crime k , let $y_{k,ct}$ and $y_{k,ct}^{pr}$ be the actual reported crime and predicted reported crime from the local linear regression for agency c during month t . We then compute the following quantity for each crime

$$\delta_{c,k} = \max_t \left(\frac{|y_{k,ct} - y_{k,ct}^{pr}|}{0.5 \times (y_{k,ct} + y_{k,ct}^{pr})} \right)$$

For each agency we compute

$$\delta_c = \max(\delta_{c,assault}; \delta_{c,burglary}; \delta_{c,larceny}; \delta_{c,robbery}; \delta_{c,motor\ vehicle\ theft})$$

We consider an agency to be an extreme outlier or to have a recorded error if $\delta_c > 1.5$. Agencies flagged as outliers are deleted from our data.

A.2 Sample selection and Cities

Table A.1: Sample construction for UCR data

Step	Description	No. Obs	No. of agency
1	RAW	2884056	22645
2	Drop agency if U.S. Park and State Police	2396724	19623
3	Drop agency if ever missing population	1924356	15349
4	Merge Crosswalk	1916040	15070
5	Keep agency if local police department	1522704	12088
6	Drop agency if not main agency	1520868	12068
7	Drop agency if reported no offense between 1985-2012	1506144	11854
8	Keep cities with population>249,999	12276	93
9	Drop agency if any outliers	9372	71
10	Merge population	9372	71
11	Drop agency if any mistakes	9372	71
12	Drop agency if any negative values for crime	8976	68
13	Drop agency if any missing values	8976	68
14	Drop agency if no crime	8976	68
15	Keep if 2007 and after	8160	68
16	Keep if balanced panel	8160	68

Table A.2: Selected cities

ORI	Department	State (1)	Min. Population (2)	Max. Population (3)	Any outliers (4)	Any #crimes <0 (5)	DOJ reform (6)
NM00101	Albuquerque Police Dpt.	New Mexico	500,955	561,560			X
CA03001	Anaheim Police Dpt.	California	333,746	353,504			
AK00101	Anchorage Police Dpt.	Alaska	277,692	301,306			
MD00202	Anne Arundel County Police Dpt.	Maryland	471,702	526,312	X		
TX22001	Arlington Police Dpt.	Texas	365,438	392,666			
GAAPD00	Atlanta Police Dpt.	Georgia	391,711	513,552			
CO00101	Aurora Police Dpt.	Colorado	259,704	320,296			
TX22701	Austin Police Dpt.	Texas	695,202	892,570			
CA01502	Bakersfield Police Dpt.	California	298,198	378,788			
MD00301	Baltimore County Police Dpt.	Maryland	785,567	831,374	X		
MDBPD00	Baltimore Police Dpt.	Maryland	618,385	638,755			X
MA01301	Boston Police Dpt.	Massachusetts	562,393	673,880			
NY01401	Buffalo Police Dpt.	New York	257,446	280,494			X
AZ00705	Chandler Police Dpt.	Arizona	236,123	265,922			
NC06001	Charlotte - Mecklenburg Police Dpt.	North Carolina	699,398	896,379			
VA02101	Chesterfield County Police Dpt.	Virginia	291,755	337,610			
ILCPD00	Chicago Police Dpt.	Illinois	2,695,598	2,857,426			X
OHCIP00	Cincinnati Police Dpt.	Ohio	296,204	333,568			X
OHCLP00	Cleveland Police Dpt.	Ohio	386,227	452,759	X		X
GA03302	Cobb County Police Dpt.	Georgia	500,622	555,827	X		
CO02101	Colorado Springs Police Dpt.	Colorado	374,112	464,113			
OHCOF00	Columbus Police Dpt.	Ohio	720,599	844,206			X
TX17802	Corpus Christi Police Dpt.	Texas	286,265	327,948			
TXDPD00	Dallas Police Dpt.	Texas	1,124,352	1,238,731			

Table A.2: Selected cities (Continued)

ORI	Department	State (1)	Min. Population (2)	Max. Population (3)	Any outliers (4)	Any #crimes <0 (5)	DOJ reform (6)
GA04402	Dekalb County Police Dpt.	Georgia	522,329	641,847			
CODPD00	Denver Police Dpt.	Colorado	568,465	699,259			
MI82349	Detroit Police Dpt.	Michigan	669,673	908,441			X
TX07102	El Paso Police Dpt.	Texas	612,374	687,193			
VA02901	Fairfax County Police Dpt.	Virginia	974,142	1,106,066			
IN00201	Fort Wayne Police	Indiana	224,820	261,642		X	
TX22012	Fort Worth Police Dpt.	Texas	640,066	842,880			
CA01005	Fresno Police Dpt.	California	465,269	524,796			
NC04102	Greensboro Police Dpt.	North Carolina	236,591	288,618	X		
GA06702	Gwinnett County Police Dpt.	Georgia	625,988	775,883		X	
NV00203	Henderson Police Dpt.	Nevada	239,906	291,584	X		
VA04301	Henrico County Division Of Police	Virginia	283,377	326,958			
HI00200	Honolulu (City And County) Police Dpt.	Hawaii	905,903	999,307	X		
TXHPD00	Houston Police Dpt.	Texas	2,031,116	2,285,709			
MD01401	Howard County Police Dpt.	Maryland	270,195	316,777			
FL01602	Jacksonville City County PD	Florida	795,822	880,557	X		
MOKPD00	Kansas City Police Dpt.	Missouri	302,696	334,250			
NV00201	Las Vegas Metro Police Dpt.	Nevada	1,315,625	1,592,178			
KY03402	Lexington Division Of Police	Kentucky	270,179	317,853			
NB05501	Lincoln Police Dpt.	Nebraska	240,511	281,138		X	
CA01941	Long Beach Police Dpt.	California	462,257	478,283			
CA01942	Los Angeles Police Dpt.	California	3,792,621	4,007,905			X
KY05680	Louisville Metro Police Dpt.	Kentucky	624,030	683,825			
TNMPD00	Memphis Police Dpt.	Tennessee	646,889	680,828			

Table A.2: Selected cities (Continued)

ORI	Department	State (1)	Min. Population (2)	Max. Population (3)	Any outliers (4)	Any #crimes <0 (5)	DOJ reform (6)
AZ00717	Mesa Police Dpt.	Arizona	439,041	478,277			
FL01306	Miami Police Dpt.	Florida	392,934	449,469	X		X
FL01300	Miami-Dade (County) Police Dpt.	Florida	1,057,396	1,208,251	X		
WIMPD00	Milwaukee Police Dpt.	Wisconsin	572,938	604,673			
MN02711	Minneapolis Police Dpt.	Minnesota	371,240	416,751	X		
AL00201	Mobile Police Dpt.	Alabama	246,171	253,842	X		
MD01604	Montgomery County Police Dpt.	Maryland	911,528	1,027,318			X
TN01901	Nashville Metro Police Dpt.	Tennessee	560,813	668,685			
NY02900	Nassau County Police Dpt.	New York	1,030,495	1,069,316	X		
DE00203	New Castle County Police Dpt.	Delaware	401,952	416,940			
LANPD00	New Orleans Police Dpt.	Louisiana	220,614	431,153			X
NY03030	New York City Police Dpt.	New York	8,165,001	8,566,917	X		
NJNPD00	Newark Police	New Jersey	277,140	281,450	X		X
CA00109	Oakland Police Dpt.	California	390,724	424,998			
OK05506	Oklahoma City Police Dpt.	Oklahoma	450,694	519,817			
NB02802	Omaha Police Dept	Nebraska	408,958	452,252			
PAPEP00	Philadelphia Police Dpt.	Pennsylvania	1,435,533	1,570,826			
AZ00723	Phoenix Police Dpt.	Arizona	1,445,632	1,597,397			
PAPPD00	Pittsburgh Police Dpt.	Pennsylvania	302,443	324,604			X
TX04306	Plano Police Dpt.	Texas	254,476	282,308			
OR02602	Portland Police Bureau	Oregon	535,763	639,601	X		X
MD01721	Prince George's County Police Dpt.	Maryland	633,458	705,672	X		X
VA07503	Prince William County Police Dpt.	Virginia	344,711	449,968			
NC09201	Raleigh Police Dpt.	North Carolina	347,755	459,399	X		

Table A.2: Selected cities (Continued)

ORI	Department	State (1)	Min. Population (2)	Max. Population (3)	Any outliers (4)	Any #crimes <0 (5)	DOJ reform (6)
CA03313	Riverside Police Dpt.	California	292,698	325,896			
CA03404	Sacramento Police Dpt.	California	460,546	495,471			
TXSPD00	San Antonio Police Dpt.	Texas	1,292,116	1,498,615			
CA03711	San Diego Police Dpt.	California	1,261,196	1,413,414			
CA03801	San Francisco Police Dpt.	California	733,799	871,155			
CA04313	San Jose Police Dpt.	California	920,548	1,041,844			
CA03019	Santa Ana Police Dpt.	California	324,528	343,433			
WASPD00	Seattle Police Dpt.	Washington	583,772	700,313			X
MOSPD00	St. Louis (City) Police Dept	Missouri	314,507	356,204			
MO09500	St. Louis County Police Dept	Missouri	387,404	412,201			
MN06209	St. Paul Police Dpt.	Minnesota	271,662	303,920		X	
CA03905	Stockton Police Dpt.	California	289,510	308,348			
NY05101	Suffolk County Police Dpt.	New York	1,308,750	1,360,739			X
FL02902	Tampa Police Dpt.	Florida	331,487	375,904		X	
OH04807	Toledo Police Dpt.	Ohio	278,366	317,401		X	
AZ01003	Tucson Police Dpt.	Arizona	520,116	547,981		X	
OK07205	Tulsa Police Dpt.	Oklahoma	375,312	399,437			
UT01800	Unified Police Dpt. of Greater Salt Lake	Utah	233,219	349,050			
VA12800	Virginia Beach Police Dpt.	Virginia	434,163	453,017			
DCMPD00	Washington Metropolitan Police Dpt.	D.C.	581,530	681,170			X
KS08703	Wichita Police Dpt.	Kansas	357,372	391,399			

Notes: This table reports the 93 police departments from cities with populations exceeding 250,000 at some point in the period analyzed (2007-2016). The fourth and fifth columns report indicate if we flag the department as an outlier or if it reported any negative crimes. The final sample does not include flagged departments. Column (6) includes cities that entered into reform agreements in pattern-or-practice policing cases ([14141 cases](#)) or are under investigation or active litigation in January 2017.

B Newsletters and Officer-involved-shooting scandal

Figure A.1: FOP Note

November 2009

Page 3

Second Vice-President's Report, *By Frank DiMaria*



It's Just a Witness Report Officer

The Lodge must receive at least 50 calls a month from our Members telling us that they were notified by either IAD or IPRA because they are being required to submit either a formal statement or a written report and that they are not accused but only a witness. The first comment from these officers sometimes will be that I am just letting the Lodge know but I don't think I need a lawyer, because I am only a witness regarding a CR investigation. The best advice I can always give to these Officers is that you are entitled to representation pursuant to Section 6.2 of our Contract which describes and details how the Department or IPRA will conduct Witness Statements involving our members whether written or oral. The affected Officer upon being notified by a supervisor or on Department computer that they are a witness in a CR investigation have the right to request to secure legal counsel. The Lodge strongly recommends that any time you must respond to this type of notification you contact the Lodge and we will provide an attorney to assist and represent you during the interview. Officers in the event

that a report is required in this investigation notify the Lodge and we will assist you in the completion of your written report. Officers don't take anything for granted in regards to these investigations. You could be a witness on Monday and then the accused on Tuesday. When notified of being a witness in a CR investigation, please contact the Lodge and we will assist you in these matters. It is your contractual right!

Parole Board Hearings

The Lodge was recently notified of the Parole Board Hearing which were conducted at 26th & California on 21 October 2009 for us to protest the possibility for parole for the convicted murderer of PO Herman Stallworth, #10965, 003rd District, End of Watch 24 May 1967.

Attending this hearing was a representative of the Department, the Police Memorial Foundation, an Assistant States Attorney (ASA), two CPD Detectives and I, representing the Lodge. The hearing which is video taped and was presided over by a member of the Illinois Prisoner Review Board who would receive our testimony and evidence. The ASA spoke to the member of the prisoner review board and retold the story of how PO Stallworth was murdered and his partner PO Ervin was shot on the night of

24 May 1967 while conducting a routine traffic stop for no other reason that they were doing their duty serving and protecting the citizens of the city of Chicago and how this coward only surrendered after he ran out of bullets in the handgun which he fired at responding police officers. The letters from the surviving widow and partner were also read and detailed how their lives had been impacted, changed forever by this violent act and that this murderer should not be released. We the representatives in attendance then all read our objections into the record asking that the Parole Board not release this convicted murderer before he complete his lawful sentence. The En Banc Hearing in this case as well as 2 other cases involving the convicted murderers of Chicago Police Officers will be heard at the State Capitol in Springfield, Illinois on 19 November 2009. We must remind those who are responsible for ensuring that these dangerous individuals never return to society.

We will travel by bus with other police officers and supporters to make our presence known and object to any possibility of parole for these convicted cop killers before their lawfully imposed sentenced is completed.

Figure A.2: Kalven Note

President's Report, by Dean C. Angelo, Sr.



Release of CR Files (Kalven Court Decision)

A recent Appellate Court decision concerning the release of CR files pursuant to Freedom of Information Act (FOIA) requests as part of *Kalven v. the City of Chicago & the Chicago Police Department* has become a major concern to the Membership; and rightfully so. When this litigation began back in 2009, the Lodge did not formally join the City or the Department in this legal challenge, thus leaving the City and the Department to challenge the FOIA requests to release CRs and RLs (Repeater Lists) on their own. Now, after the decision has been heard and appealed, the Lodge has no legal standing or litigant position to file any additional actions.

So what does this mean to the Membership? In compliance with the court decision, lists of past complaints against Chicago Police Officers will now be released to the public for review. We initially spoke about this case at our first General Meeting in April and mentioned then that this decision would wind up being something that every working officer would have an extensive interest in. We contacted the Lodge's attorneys and requested their take on the impact. (Further information that details the attorney's input is available on page six of this Newsletter).

More History

In December 2009, (Attorney) Kalven filed suit against the City (and the CPD) under the Freedom of Information Act (FOIA), to obtain Complaint Register (CR) files as well as Repeater Lists (RL). Initially, the Circuit Court held that Kalven was entitled to the RLs, but not to the CRs. After an appeal, the Court ruled Kalven was entitled to both the RLs and CRs, (subject to complainant and witness redaction).

The intention of the Freedom of Information Act (as argued in Kalven), was to allow the public to review the actions of public bodies to ensure full disclosure of public employees and make them available for review; such as with Chicago Police Officers. An individual working for a private employer would not face this level of scrutiny. As we are all well aware, a CR Investigation of a Police Officer does not establish any wrongdoing on the part of the Officer; it is a fact that most allegations are either Not Sustained, Unfounded or, Exonerated.

An additional concern of ours is the impact this type of disclosure might have on the reputation of the individual Police Officers and quite frankly, the Chicago Police Department as well. A Police Officer might suffer undue scrutiny from family members, friends, neighbors and even co-workers, when this information is released. Furthermore, an additional risk now exists that the public (including some criminal defense attorneys) can start to access this information through a Freedom of Information Act request in a much more exaggerated manner.

Lodge Position

The Lodge emphatically believes the Court erred in its ruling. Consequently, we believe the statute should be amended to exempt both CRs and RLs from public view and production. The Lodge sees this decision as unfair prejudice and that such disclosures may have a negative impact on individual Police Officers. There also exists a likelihood that such public exposure will discourage many intelligent, skilled and dedicated individuals from continuing, or even beginning a career in public service. When is enough, truly enough?

A Final Update on Kalven v. City

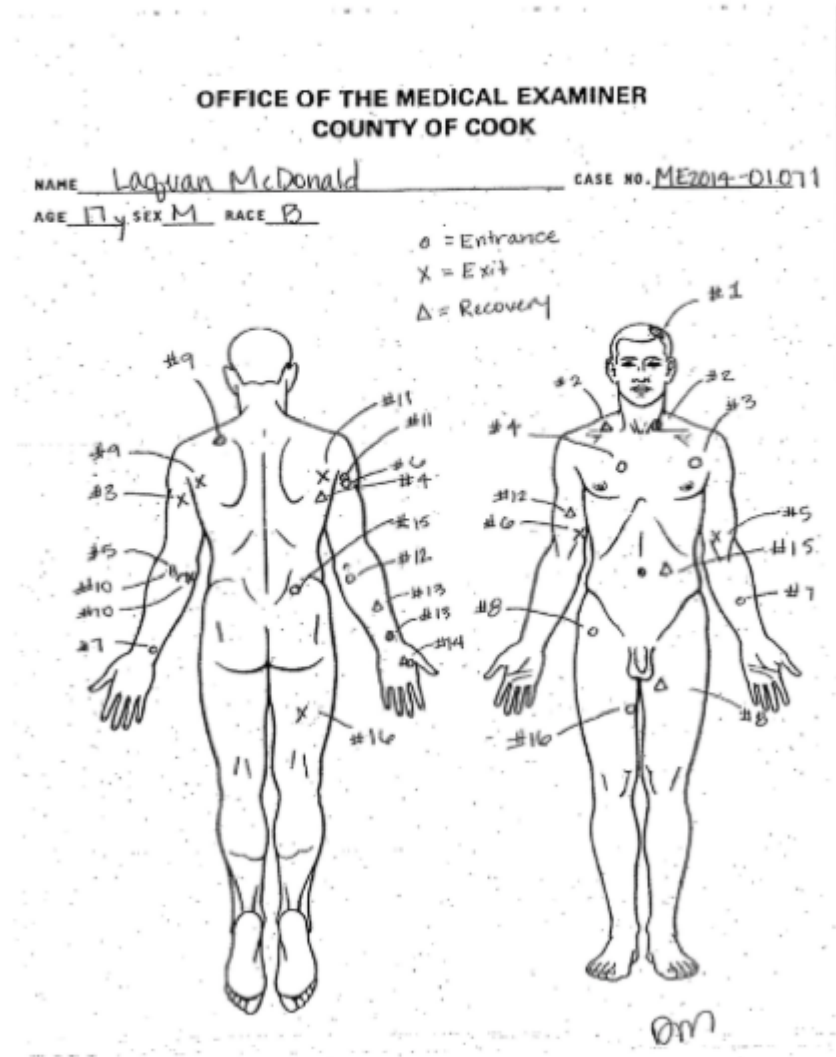
Although we cannot alter the Court's decision concerning the release of the CR and RL lists, we might have one option available to possibly challenge some of the names on those lists. On 28 July 14, the Lodge filed an Emergency Motion requesting a Temporary Restraining Order

(Continued on page 2)

Figure A.3: Officer-involved-shooting scandal

The New York Times

Video of Chicago Police Shooting a Teenager Is Ordered Released



An autopsy drawing of Laquan McDonald, 17, who was fatally shot by a Chicago police officer in October 2014. According to a report from the Cook County medical examiner's office, he was shot 16 times. Cook County Medical Examiner

C Additional tables and figures

Figure A.4: CPD Areas, Districts and Beats

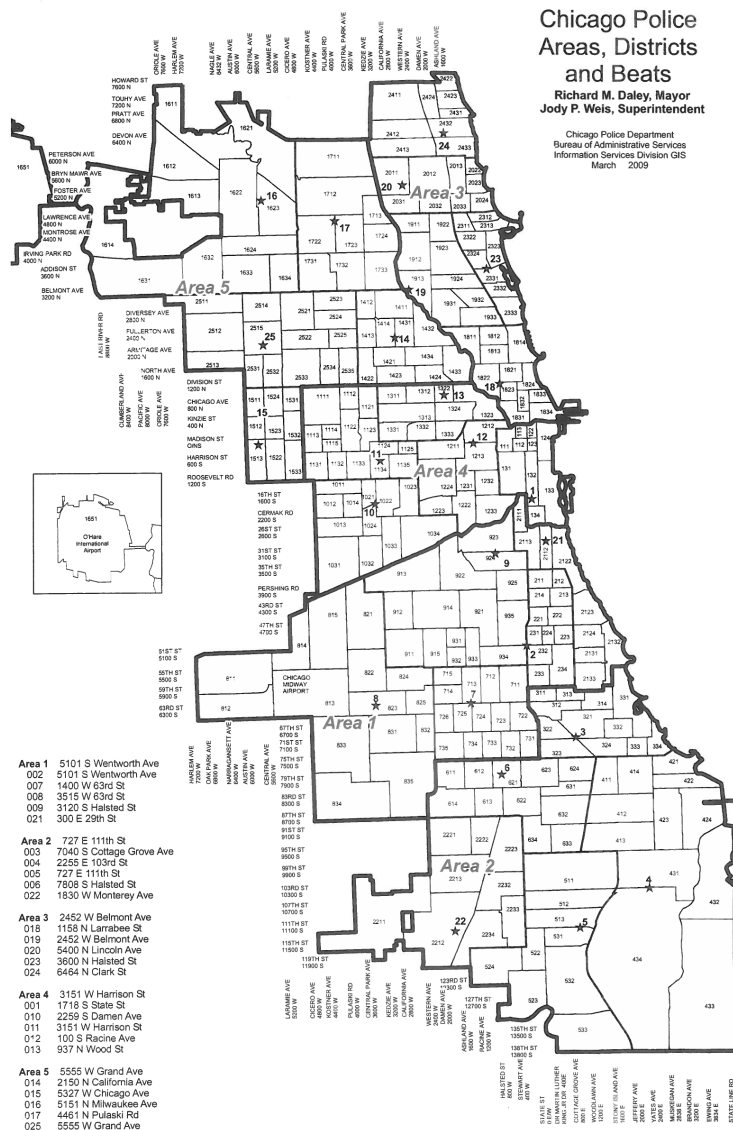
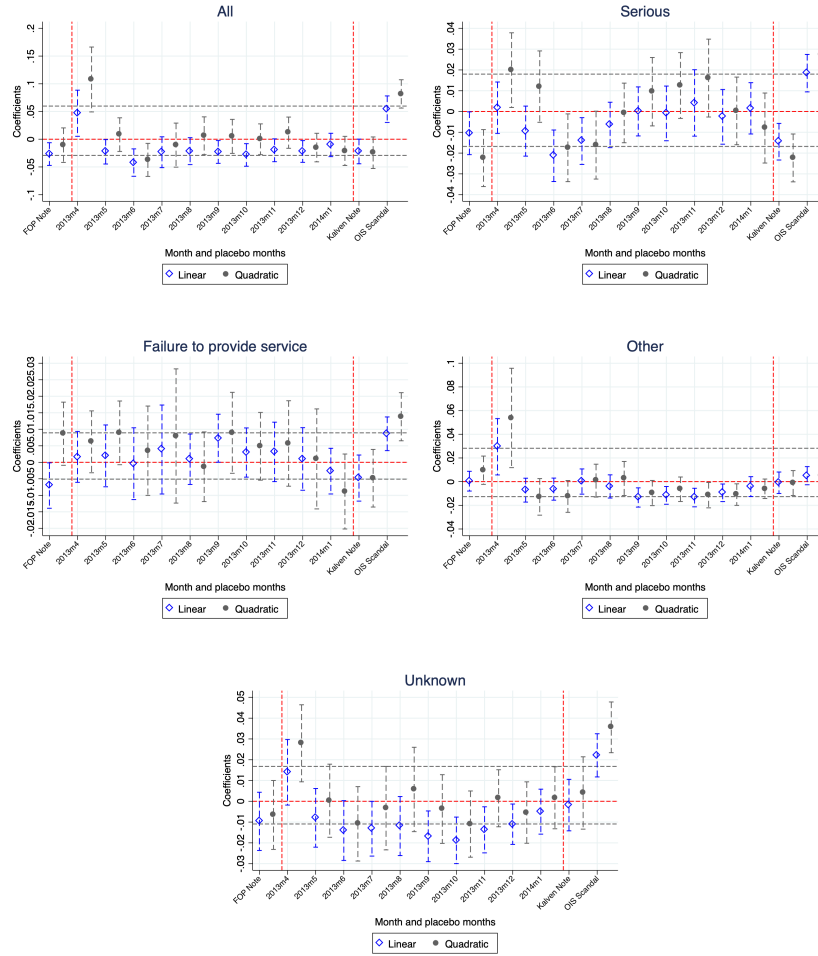
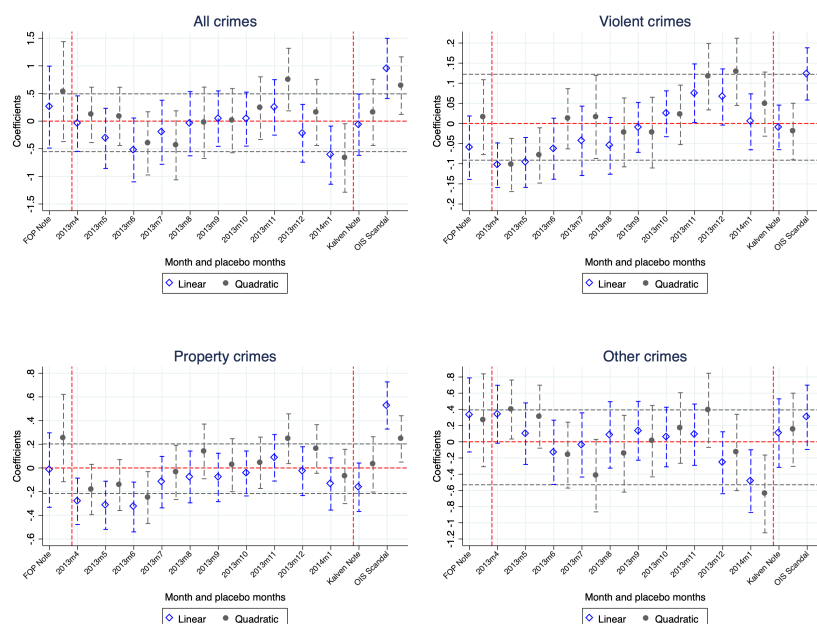


Figure A.5: Placebo estimates of the effect of oversight on complaints



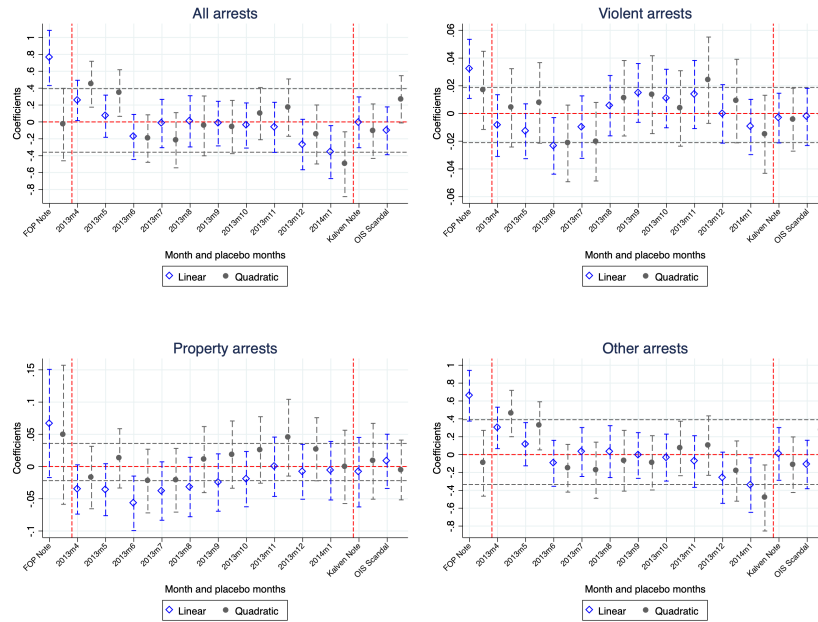
Notes: Each placebo estimate first assigns a window around a false event date from April 2013 to January 2014, and then uses an RD to estimate the effect of oversight on complaints. The placebo estimates are not independent of each other, as the samples overlap. Standard errors are clustered at the police beat level and we report the 95% confidence interval of each point estimates. The dotted horizontal grey lines correspond to the 10th and 90th percentiles of the placebos of the quadratic specification.

Figure A.6: Placebo estimates of the effect of oversight on crimes



Notes: Each placebo estimate first assigns a window around a false event date from April 2013 to January 2014, and then uses an RD to estimate the effect of oversight on crimes. The placebo estimates are not independent of each other, as the samples overlap. Standard errors are clustered at the police beat level and we report the 95% confidence interval of each point estimates. The dotted horizontal grey lines correspond to the 10th and 90th percentiles of the placebos of the quadratic specification.

Figure A.7: Placebo estimates of the effect of oversight on arrests



Notes: Each placebo estimate first assigns a window around a false event date from April 2013 to January 2014, and then uses an RD to estimate the effect of oversight on arrests. The placebo estimates are not independent of each other, as the samples overlap. Standard errors are clustered at the police beat level and we report the 95% confidence interval of each point estimates. The dotted horizontal grey lines correspond to the 10th and 90th percentiles of the placebos of the quadratic specification.

Table A.3: Complaints response to monitoring-Donut RD dropping month of the event

	All (1)	All (2)	Serious (3)	Serious (4)	FPS (5)	FPS (6)	Other (7)	Other (8)	Unknown (9)	Unknown (10)
Panel A, FOP Note										
Event	-0.030** (0.012)	-0.058** (0.026)	-0.008 (0.006)	-0.050*** (0.012)	-0.012*** (0.004)	-0.003 (0.007)	0.001 (0.005)	0.012 (0.009)	-0.010 (0.009)	-0.017 (0.013)
Mean pre-event	0.259	0.259	0.088	0.088	0.034	0.034	0.040	0.040	0.098	0.098
Panel B, Kalven Note										
Event	-0.040*** (0.014)	-0.070*** (0.019)	-0.020*** (0.006)	-0.047*** (0.009)	-0.006 (0.005)	-0.006 (0.007)	-0.000 (0.006)	-0.004 (0.007)	-0.014** (0.006)	-0.013 (0.011)
Mean pre-event	0.197	0.197	0.063	0.063	0.025	0.025	0.033	0.033	0.076	0.076
Panel C, OIS Scandal										
Event	0.065*** (0.013)	0.133*** (0.016)	0.023*** (0.005)	0.044*** (0.008)	0.009*** (0.003)	0.014*** (0.004)	0.006 (0.004)	0.012* (0.007)	0.028*** (0.006)	0.064*** (0.009)
Mean pre-event	0.146	0.146	0.043	0.043	0.018	0.018	0.029	0.029	0.055	0.055
No of obs.	9108	9108	9108	9108	9108	9108	9108	9108	9108	9108
No of beats	276	276	276	276	276	276	276	276	276	276
Specification	Linear	Quadr.	Linear	Quadr.	Linear	Quadr.	Linear	Quadr.	Linear	Quadr.

Notes: This Table presents the effect of oversights on the the number of complaints per 10,000 residents in Chicago from January 2007 to May 2016 (see Section 2). We account for seasonal patterns in outcomes by removing the month effect and report the residuals of this regression. We report the mean of the dependent variable at pre-event and RD estimates. Standard errors are clustered at the police beat level and are reported in parentheses. *p-value < 0.10, **p-value < 0.05, *** p-value < 0.01.

Table A.4: Crimes response to monitoring-Donut RD dropping month of the event

	All Crimes (1)	All Crimes (2)	Violent Crimes (3)	Violent Crimes (4)	Property Crimes (5)	Property Crimes (6)	Other Crimes (7)	Other Crimes (8)
Panel A, FOP Note								
Event	0.457 (0.371)	-0.209 (0.696)	-0.055 (0.046)	-0.046 (0.066)	0.005 (0.160)	0.307 (0.263)	0.508** (0.233)	-0.470 (0.460)
Mean pre-event	15.608	15.608	1.285	1.285	4.693	4.693	9.630	9.630
Panel B, Kalven Note								
Event	-0.082 (0.278)	-0.072 (0.324)	0.033 (0.038)	0.067 (0.062)	-0.148 (0.108)	0.163 (0.144)	0.033 (0.214)	-0.302 (0.258)
Mean pre-event	13.056	13.056	1.088	1.088	4.034	4.034	7.933	7.933
Panel C, OIS Scandal								
Event	1.244*** (0.308)	1.039*** (0.318)	0.181*** (0.038)	0.117* (0.067)	0.740*** (0.107)	0.461*** (0.138)	0.324 (0.233)	0.461* (0.263)
Mean pre-event	11.805	11.805	0.991	0.991	3.456	3.456	7.357	7.357
No of obs.	9108	9108	9108	9108	9108	9108	9108	9108
No of beats	276	276	276	276	276	276	276	276
Specification	Linear	Quadr.	Linear	Quadr.	Linear	Quadr.	Linear	Quadr.

Notes: This Table presents the effect of oversights on the the number of crimes per 10,000 residents in Chicago from January 2007 to May 2016 (see Section 2). We account for seasonal patterns in outcomes by removing the month effect and report the residuals of this regression. We report the mean of the dependent variable at pre-event and RD estimates. Standard errors are clustered at the police beat level and are reported in parentheses. *p-value < 0.10, **p-value < 0.05, *** p-value < 0.01.

Table A.5: Arrests response to monitoring-Donut RD dropping month of the event

	All Arrests (1)	All Arrests (2)	Violent Arrests (3)	Violent Arrests (4)	Property Arrests (5)	Property Arrests (6)	Other Arrests (7)	Other Arrests (8)
Panel A, FOP Note								
Event	0.949*** (0.172)	-0.512 (0.330)	0.031** (0.014)	-0.014 (0.024)	0.062 (0.042)	0.019 (0.068)	0.856*** (0.151)	-0.517* (0.290)
Mean pre-event	4.607	4.607	0.229	0.229	0.467	0.467	3.911	3.911
Panel B, Kalven Note								
Event	-0.015 (0.153)	-0.370** (0.188)	0.005 (0.013)	0.016 (0.023)	-0.018 (0.032)	0.024 (0.042)	-0.003 (0.150)	-0.410** (0.180)
Mean pre-event	3.952	3.952	0.194	0.194	0.381	0.381	3.377	3.377
Panel C, OIS Scandal								
Event	-0.268 (0.164)	0.361* (0.189)	0.002 (0.013)	0.012 (0.021)	0.021 (0.025)	-0.013 (0.040)	-0.291* (0.157)	0.361** (0.178)
Mean pre-event	3.606	3.606	0.182	0.182	0.352	0.352	3.073	3.073
No of obs.	9108	9108	9108	9108	9108	9108	9108	9108
No of beats	276	276	276	276	276	276	276	276
Specification	Linear	Quadr.	Linear	Quadr.	Linear	Quadr.	Linear	Quadr.

Notes: This Table presents the effect of oversights on the the number of arrests per 10,000 residents in Chicago from January 2007 to May 2016 (see Section 2). We account for seasonal patterns in outcomes by removing the month effect and report the residuals of this regression. We report the mean of the dependent variable at pre-event and RD estimates. Standard errors are clustered at the police beat level and are reported in parentheses. *p-value < 0.10, **p-value < 0.05, *** p-value < 0.01.

Table A.6: Complaints response to monitoring-Donut RD dropping months within 0 and -1

	All (1)	All (2)	Serious (3)	Serious (4)	FPS (5)	FPS (6)	Other (7)	Other (8)	Unknown (9)	Unknown (10)
Panel A, FOP Note										
Event	-0.033*** (0.013)	-0.057** (0.027)	-0.009 (0.006)	-0.051*** (0.012)	-0.014*** (0.004)	-0.003 (0.007)	0.001 (0.005)	0.015 (0.010)	-0.011 (0.009)	-0.017 (0.013)
Mean pre-event	0.261	0.261	0.088	0.088	0.034	0.034	0.040	0.040	0.099	0.099
Panel B, Kalven Note										
Event	-0.041*** (0.014)	-0.075*** (0.020)	-0.021*** (0.006)	-0.052*** (0.009)	-0.006 (0.005)	-0.007 (0.008)	-0.000 (0.006)	-0.004 (0.007)	-0.013** (0.006)	-0.012 (0.011)
Mean pre-event	0.197	0.197	0.064	0.064	0.025	0.025	0.033	0.033	0.076	0.076
Panel C, OIS Scandal										
Event	0.069*** (0.013)	0.145*** (0.017)	0.023*** (0.005)	0.045*** (0.008)	0.010*** (0.003)	0.016*** (0.004)	0.007 (0.005)	0.014** (0.007)	0.030*** (0.006)	0.070*** (0.010)
Mean pre-event	0.146	0.146	0.044	0.044	0.018	0.018	0.029	0.029	0.055	0.055
No of obs.	8556	8556	8556	8556	8556	8556	8556	8556	8556	8556
No of beats	276	276	276	276	276	276	276	276	276	276
Specification	Linear	Quadr.	Linear	Quadr.	Linear	Quadr.	Linear	Quadr.	Linear	Quadr.

Notes: This Table presents the effect of oversights on the the number of complaints per 10,000 residents in Chicago from January 2007 to May 2016 (see Section 2). We account for seasonal patterns in outcomes by removing the month effect and report the residuals of this regression. We report the mean of the dependent variable at pre-event and RD estimates. Standard errors are clustered at the police beat level and are reported in parentheses. *p-value < 0.10, **p-value < 0.05, *** p-value < 0.01.

Table A.7: Crimes response to monitoring-Donut RD dropping dropping months within 0 and -1

	All Crimes (1)	All Crimes (2)	Violent Crimes (3)	Violent Crimes (4)	Property Crimes (5)	Property Crimes (6)	Other Crimes (7)	Other Crimes (8)
Panel A, FOP Note								
Event	0.458 (0.376)	0.009 (0.703)	-0.071 (0.047)	-0.057 (0.068)	0.004 (0.163)	0.387 (0.270)	0.524** (0.236)	-0.321 (0.462)
Mean pre-event	15.642	15.642	1.288	1.288	4.692	4.692	9.662	9.662
Panel B, Kalven Note								
Event	-0.086 (0.280)	-0.003 (0.331)	0.038 (0.039)	0.082 (0.063)	-0.159 (0.108)	0.183 (0.148)	0.035 (0.217)	-0.268 (0.266)
Mean pre-event	13.048	13.048	1.085	1.085	4.035	4.035	7.928	7.928
Panel C, OIS Scandal								
Event	1.280*** (0.312)	1.095*** (0.326)	0.183*** (0.038)	0.115* (0.067)	0.753*** (0.110)	0.479*** (0.141)	0.344 (0.236)	0.500* (0.270)
Mean pre-event	11.772	11.772	0.981	0.981	3.443	3.443	7.348	7.348
No of obs.	8556	8556	8556	8556	8556	8556	8556	8556
No of beats	276	276	276	276	276	276	276	276
Specification	Linear	Quadr.	Linear	Quadr.	Linear	Quadr.	Linear	Quadr.

Notes: This Table presents the effect of oversights on the the number of crimes per 10,000 residents in Chicago from January 2007 to May 2016 (see Section 2). We account for seasonal patterns in outcomes by removing the month effect and report the residuals of this regression. We report the mean of the dependent variable at pre-event and RD estimates. Standard errors are clustered at the police beat level and are reported in parentheses. *p-value < 0.10, **p-value < 0.05, *** p-value < 0.01.

Table A.8: Arrests response to monitoring-Donut RD dropping months within 0 and -1

	All Arrests (1)	All Arrests (2)	Violent Arrests (3)	Violent Arrests (4)	Property Arrests (5)	Property Arrests (6)	Other Arrests (7)	Other Arrests (8)
Panel A, FOP Note								
Event	1.012*** (0.173)	-0.469 (0.332)	0.032** (0.014)	-0.014 (0.025)	0.063 (0.042)	0.019 (0.069)	0.916*** (0.152)	-0.474 (0.292)
Mean pre-event	4.615	4.615	0.230	0.230	0.466	0.466	3.919	3.919
Panel B, Kalven Note								
Event	0.003 (0.157)	-0.334* (0.198)	0.005 (0.013)	0.015 (0.023)	-0.019 (0.032)	0.020 (0.043)	0.017 (0.155)	-0.369* (0.191)
Mean pre-event	3.947	3.947	0.194	0.194	0.380	0.380	3.373	3.373
Panel C, OIS Scandal								
Event	-0.252 (0.166)	0.393** (0.197)	0.001 (0.013)	0.012 (0.021)	0.020 (0.025)	-0.011 (0.041)	-0.273* (0.160)	0.392** (0.185)
Mean pre-event	3.606	3.606	0.182	0.182	0.350	0.350	3.074	3.074
No of obs.	8556	8556	8556	8556	8556	8556	8556	8556
No of beats	276	276	276	276	276	276	276	276
Specification	Linear	Quadr.	Linear	Quadr.	Linear	Quadr.	Linear	Quadr.

Notes: This Table presents the effect of oversights on the the number of arrests per 10,000 residents in Chicago from January 2007 to May 2016 (see Section 2). We account for seasonal patterns in outcomes by removing the month effect and report the residuals of this regression. We report the mean of the dependent variable at pre-event and RD estimates. Standard errors are clustered at the police beat level and are reported in parentheses. *p-value < 0.10, **p-value < 0.05, *** p-value < 0.01.

Table A.9: Complaints response to monitoring-Donut RD dropping months within 0 and -2

	All (1)	All (2)	Serious (3)	Serious (4)	FPS (5)	FPS (6)	Other (7)	Other (8)	Unknown (9)	Unknown (10)
Panel A, FOP Note										
Event	-0.036*** (0.013)	-0.058** (0.028)	-0.009 (0.006)	-0.053*** (0.012)	-0.015*** (0.004)	-0.002 (0.008)	-0.000 (0.005)	0.015 (0.010)	-0.013 (0.009)	-0.019 (0.014)
Mean pre-event	0.263	0.263	0.089	0.089	0.034	0.034	0.041	0.041	0.100	0.100
Panel B, Kalven Note										
Event	-0.041*** (0.014)	-0.078*** (0.021)	-0.020*** (0.006)	-0.051*** (0.009)	-0.006 (0.005)	-0.008 (0.008)	-0.001 (0.006)	-0.006 (0.008)	-0.014** (0.006)	-0.015 (0.011)
Mean pre-event	0.198	0.198	0.064	0.064	0.025	0.025	0.033	0.033	0.076	0.076
Panel C, OIS Scandal										
Event	0.076*** (0.013)	0.164*** (0.018)	0.026*** (0.005)	0.051*** (0.008)	0.010*** (0.003)	0.019*** (0.005)	0.007 (0.005)	0.016** (0.007)	0.032*** (0.006)	0.078*** (0.010)
Mean pre-event	0.148	0.148	0.044	0.044	0.018	0.018	0.029	0.029	0.056	0.056
No of obs.	8004	8004	8004	8004	8004	8004	8004	8004	8004	8004
No of beats	276	276	276	276	276	276	276	276	276	276
Specification	Linear	Quadr.	Linear	Quadr.	Linear	Quadr.	Linear	Quadr.	Linear	Quadr.

Notes: This Table presents the effect of oversights on the the number of complaints per 10,000 residents in Chicago from January 2007 to May 2016 (see Section 2). We account for seasonal patterns in outcomes by removing the month effect and report the residuals of this regression. We report the mean of the dependent variable at pre-event and RD estimates. Standard errors are clustered at the police beat level and are reported in parentheses. *p-value < 0.10, **p-value < 0.05, *** p-value < 0.01.

Table A.10: Crimes response to monitoring-Donut RD dropping months within 0 and -2

	All Crimes (1)	All Crimes (2)	Violent Crimes (3)	Violent Crimes (4)	Property Crimes (5)	Property Crimes (6)	Other Crimes (7)	Other Crimes (8)
Panel A, FOP Note								
Event	0.399 (0.383)	0.120 (0.715)	-0.088* (0.047)	-0.076 (0.068)	-0.004 (0.165)	0.465* (0.275)	0.492** (0.242)	-0.269 (0.470)
Mean pre-event	15.671	15.671	1.288	1.288	4.688	4.688	9.695	9.695
Panel B, Kalven Note								
Event	-0.104 (0.281)	0.048 (0.341)	0.036 (0.039)	0.083 (0.065)	-0.168 (0.110)	0.216 (0.153)	0.028 (0.220)	-0.251 (0.275)
Mean pre-event	13.058	13.058	1.084	1.084	4.046	4.046	7.928	7.928
Panel C, OIS Scandal								
Event	1.326*** (0.315)	1.210*** (0.334)	0.186*** (0.038)	0.128* (0.068)	0.776*** (0.112)	0.523*** (0.145)	0.363 (0.238)	0.559** (0.276)
Mean pre-event	11.758	11.758	0.973	0.973	3.438	3.438	7.346	7.346
No of obs.	8004	8004	8004	8004	8004	8004	8004	8004
No of beats	276	276	276	276	276	276	276	276
Specification	Linear	Quadr.	Linear	Quadr.	Linear	Quadr.	Linear	Quadr.

Notes: This Table presents the effect of oversights on the the number of crimes per 10,000 residents in Chicago from January 2007 to May 2016 (see Section 2). We account for seasonal patterns in outcomes by removing the month effect and report the residuals of this regression. We report the mean of the dependent variable at pre-event and RD estimates. Standard errors are clustered at the police beat level and are reported in parentheses. *p-value < 0.10, **p-value < 0.05, *** p-value < 0.01.

Table A.11: Arrests response to monitoring-Donut RD dropping months within 0 and -2

	All Arrests (1)	All Arrests (2)	Violent Arrests (3)	Violent Arrests (4)	Property Arrests (5)	Property Arrests (6)	Other Arrests (7)	Other Arrests (8)
Panel A, FOP Note								
Event	1.047*** (0.175)	-0.495 (0.339)	0.034** (0.014)	-0.011 (0.025)	0.065 (0.042)	0.019 (0.070)	0.948*** (0.155)	-0.502* (0.298)
Mean pre-event	4.625	4.625	0.230	0.230	0.466	0.466	3.929	3.929
Panel B, Kalven Note								
Event	0.026 (0.160)	-0.273 (0.208)	0.006 (0.013)	0.019 (0.023)	-0.018 (0.033)	0.023 (0.045)	0.038 (0.157)	-0.316 (0.200)
Mean pre-event	3.949	3.949	0.193	0.193	0.381	0.381	3.375	3.375
Panel C, OIS Scandal								
Event	-0.246 (0.170)	0.391* (0.207)	-0.000 (0.013)	0.009 (0.022)	0.024 (0.025)	-0.002 (0.042)	-0.270* (0.163)	0.385** (0.194)
Mean pre-event	3.615	3.615	0.180	0.180	0.350	0.350	3.084	3.084
No of obs.	8004	8004	8004	8004	8004	8004	8004	8004
No of beats	276	276	276	276	276	276	276	276
Specification	Linear	Quadr.	Linear	Quadr.	Linear	Quadr.	Linear	Quadr.

Notes: This Table presents the effect of oversights on the the number of arrests per 10,000 residents in Chicago from January 2007 to May 2016 (see Section 2). We account for seasonal patterns in outcomes by removing the month effect and report the residuals of this regression. We report the mean of the dependent variable at pre-event and RD estimates. Standard errors are clustered at the police beat level and are reported in parentheses. *p-value < 0.10, **p-value < 0.05, *** p-value < 0.01.

Table A.12: Effect of monitoring on crime using alternative samples

	Sample 1			Sample 2			Sample 3		
	FOP (1)	Kalven (2)	OIS (3)	FOP (4)	Kalven (5)	OIS (6)	FOP (7)	Kalven (8)	OIS (9)
Violent	-8.944 (15.487)	-5.527 (12.127)	10.717 (19.575)	-11.187 (10.641)	-7.052 (12.227)	10.032 (12.404)	-6.816 (7.128)	-7.427 (7.764)	7.317 (13.176)
Mean Chicago _{t<t₀}	98.854	75.806	71.793	98.854	75.806	71.793	98.854	75.806	71.793
Assault	-1.949 (16.12)	-3.744 (9.956)	5.802 (23.985)	-3.479 (8.264)	-4.323 (10.337)	6.434 (13.052)	-4.68 (7.144)	-1.57 (10.569)	6.728 (32.414)
Mean Chicago _{t<t₀}	50.206	37.096	38.673	50.206	37.096	38.673	50.206	37.096	38.673
Murder	0.02 (0.222)	0.034 (0.353)	0.671*** (0.205)	0.004 (0.191)	0.043 (0.342)	0.578*** (0.162)	0.016 (0.153)	0.105 (0.201)	0.664*** (0.222)
Mean Chicago _{t<t₀}	1.397	1.353	1.341	1.397	1.353	1.341	1.397	1.353	1.341
Robbery	-1.876 (3.067)	-6.645** (3.248)	8.83*** (2.602)	-1.576 (3.239)	-8.63** (2.989)	7.899*** (2.015)	-0.807 (3.204)	-3.74 (2.189)	10.408*** (3.479)
Mean Chicago _{t<t₀}	47.251	37.358	31.779	47.251	37.358	31.779	47.251	37.358	31.779
Property	15.582 (23.246)	-3.346 (26.974)	28.26 (25.891)	13.357 (20.448)	0.273 (28.08)	20.94 (19.875)	23.646 (28.036)	0.234 (23.274)	25.452 (18.095)
Mean Chicago _{t<t₀}	373.701	311.613	269.479	373.701	311.613	269.479	373.701	311.613	269.479
Burglary	-0.151 (8.486)	-7.263 (8.449)	3.219 (7.312)	1.16 (7.997)	-5.451 (7.45)	3.658 (6.147)	2.637 (8.146)	-3.485 (5.946)	4.88 (4.813)
Mean Chicago _{t<t₀}	75.79	60.172	46.479	75.79	60.172	46.479	75.79	60.172	46.479
Larceny	25.117 (38.594)	-9.236 (19.001)	14.346 (16.52)	-15.422 (33.23)	-7.664 (19.348)	8.229 (16.905)	8.07 (18.821)	-7.559 (15.237)	12.226 (15.245)
Mean Chicago _{t<t₀}	245.693	208.228	189.426	245.693	208.228	189.426	245.693	208.228	189.426
Motor Vehicle Theft	15.607** (7.222)	-13.85 (10.657)	0.972 (10.847)	12.865** (5.966)	-7.637 (5.59)	0.136 (11.367)	12.941* (6.826)	3.762 (16.523)	5.341 (6.235)
Mean Chicago _{t<t₀}	52.219	43.213	33.574	52.219	43.213	33.574	52.219	43.213	33.574
No. of Cities	68	68	68	56	56	56	33	33	33
No. of Observation	3264	3264	3264	2688	2688	2688	1584	1584	1584
Population \geq 500,00	No	No	No	No	No	No	Yes	Yes	Yes
DOJ Investigated Excluded	No	No	No	Yes	Yes	Yes	No	No	No

Notes: This Table presents the effect of oversight on crimes per 100,000 capita (see Section 2). We account for seasonal patterns in crime outcomes by removing the city-month effect and report the residuals of this regression. Standard errors for the generalized synthetic controls methods (GSC) are based on parametric bootstraps (blocked at the agency level) of 10,000 times. *p-value < 0.10, **p-value < 0.05, *** p-value < 0.01.

Table A.13: Effect of monitoring on crime using Synthetic Control Method

	Violent (1)	Assault (2)	Murder (3)	Robbery (4)	Property (5)	Burglary (6)	Larceny (7)	Auto Theft (8)
Panel A, FOP Note								
ATT	-11.386	-4.457	0.183	5.421	31.658	2.664	2.304	15.518
5th percentile	-11.382	-14.794	-0.205	-4.893	-61.398	-14.358	-22.126	-5.297
95th percentile	13.796	15.096	0.375	4.733	68.042	19.059	62.075	10.337
p-value	0.458	0.833	0.250	0.083	0.333	0.875	0.875	0.042
Panel B, Kalven Note								
ATT	-8.705	-3.261	0.061	-7.323	-7.915	-6.757	-10.058	-5.733
5th percentile	-12.689	-14.120	-0.200	-3.847	-32.141	-10.344	-25.330	-6.730
95th percentile	23.465	25.668	0.335	3.913	71.624	10.528	43.022	11.009
p-value	0.667	0.792	0.375	0.083	0.833	0.542	0.875	0.375
Panel C, OIS Scandal								
ATT	11.545	2.316	0.500	6.603	18.968	5.657	9.910	1.671
5th percentile	-12.754	-14.661	-0.285	-3.971	-18.370	-8.901	-16.435	-6.034
95th percentile	27.805	25.816	0.488	4.102	35.471	6.878	32.296	16.195
p-value	0.250	0.708	0.083	0.083	0.417	0.625	0.583	0.750
No. of Cities	25	25	25	25	25	25	25	25
No. of Observation	1200	1200	1200	1200	1200	1200	1200	1200

Notes: This Table presents the effect of oversight on crimes per 100,000 capita (see Section 2) using synthetic control methods. We account for seasonal patterns in crime outcomes by removing the month effect and use the residuals of this regression as the outcome. We conduct a two-tailed test of the null hypothesis of no effect in our treatment by comparing the observed estimate for Chicago to the empirical distribution of the placebo estimates. We report the 5th-95th percentile of the placebo estimates.